

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:42:39 ; Search time 14 Seconds  
(without alignments)  
2044.190 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 690  
Sequence: 1 MRAAPLLIARAASLSIGFLF.....EPVQEGAPPPAAHHHHH 690

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	292	42.3	1255	1	ERR2_HUMAN
2	127	18.4	144	1	CSF2_HUMAN
3	67	9.7	1254	1	ERR2_MESAU
4	52	7.5	1257	1	ERR2_RAT
5	35	5.1	386	1	PPAP_HUMAN
6	22	3.2	144	1	CSF2_SHEEP
7	20	2.9	245	1	ERR2_MOUSE
8	13	1.9	140	1	CSF2_MOUSE
9	11	1.6	144	1	CSF2_CANFA
10	11	1.6	144	1	CSF2_CEREL
11	11	1.6	144	1	CSF2_PIG
12	10	1.4	143	1	CSF2_BOVIN
13	9	1.3	530	1	ZIC2_MOUSE
14	9	1.3	532	1	ZIC2_HUMAN
15	9	1.3	703	1	EGFR_CHICK
16	9	1.3	1210	1	EGFR_HUMAN
17	9	1.3	1210	1	EGFR_MOUSE
18	8	1.2	127	1	CSF2_RAT
19	8	1.2	265	1	UCR1_SOLITU
20	8	1.2	325	1	MOD1_MYCHO
21	8	1.2	325	1	MOD1_MYCHO
22	8	1.2	355	1	MONA_ARTNI
23	8	1.2	428	1	FXB2_MOUSE
24	8	1.2	463	1	VY30_MYCTU
25	8	1.2	547	1	CITR_KLEBN
26	8	1.2	715	1	CLPB_MYCPN
27	8	1.2	825	1	ICP0_HSV2H
28	8	1.2	1848	1	CCNA_DROME
29	7	1.0	62	1	SECE_SULSO
30	7	1.0	68	1	RI29_ARCTU
31	7	1.0	104	1	YIF4_ARCTU
32	7	1.0	111	1	MADE_HUMAN
33	7	1.0	125	1	CALR_CHICK

34	7	1.0	136	1	CALL_ONCKE
35	7	1.0	138	1	CAL_CHICK
36	7	1.0	141	1	CSF2_MOUSE
37	7	1.0	144	1	CSF2_FELCA
38	7	1.0	182	1	NCC2_RHME
39	7	1.0	184	1	RUS_THEMA
40	7	1.0	186	1	RK12_NICSY
41	7	1.0	186	1	RK12_TOBAC
42	7	1.0	196	1	UREE_ALCEU
43	7	1.0	196	1	GIDB_AQUAE
44	7	1.0	204	1	GIDB_COXBU
45	7	1.0	204	1	HANI_SHEEP

## ALIGNMENTS

RESULT 1	ERR2_HUMAN	STANDARD:	PRT: 1255 AA.
ID	ERR2_HUMAN		
AC	P04626:		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)		
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).		
DE	surface receptor HER2 (MLN 19).		
GN	ERRB2 OR HER2 OR NGL OR NEU.		
OC	Human sapliens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID:9606;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-86118663; PubMed-3003577;		
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,		
RA	Saito T., Toyoshima K.;		
RT	"Similarity of protein encoded by the human c-erbB-2 gene to		
RT	epidermal growth factor receptor.";		
RL	Nature 319:230-234(1986).		
[2]			
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-86070181; PubMed-2999974;		
RA	Consensus L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,		
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,		
RA	Francake U., Levinson A., Ullrich A.;		
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor		
RT	shares chromosomal location with neu oncogene.";		
RL	Science 230:1132-1139(1985).		
[3]			
RP	SEQUENCE OF 737-1031 FROM N.A.		
RX	MEDLINE-86016729; PubMed-2959567;		
RA	Sema K., Kamata N., Toyoshima K., Yamamoto T.;		
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the		
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a		
RT	human salivary gland adenocarcinoma.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).		
[4]			
RP	VARIANTS VAL-654 AND VAL-655.		
RX	MEDLINE-93194196; PubMed-8095488;		
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;		
RT	"Characterization of a new allele of the human ERBB2 gene by allele-		
RT	specific competition hybridization.";		
RL	Genomics 15:426-429(1993).		
-1-	FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,		
CC	ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GFI3 IS A		
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-		
CC	ALPHA AND AMPHIREGULIN.		
-1-	CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein		
CC	tyrosine phosphate.		
CC	-1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS		
CC	(POTENTIAL).		
-1-	SUBCELLULAR LOCATION: Type I membrane protein.		

Query Match  
Best Local Similarity 42.3%; Score 292; DB 1; Length 1255;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FT	DISULFID	563	576	BY SIMILARITY.
FT	DISULFID	567	584	BY SIMILARITY.
FT	DISULFID	587	596	BY SIMILARITY.
FT	DISULFID	600	623	BY SIMILARITY.
FT	DISULFID	626	634	BY SIMILARITY.
FT	DISULFID	630	642	BY SIMILARITY.
FT	MOD_RES	1139	1139	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1248	1248	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	CARBOHYD	68	68	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	124	124	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	187	187	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	530	530	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	571	571	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	629	629	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	654	654	I -> V.
FT	VARIANT	655	655	/FTID=VAR_004077.
FT	VARIANT	655	655	I -> V.
FT	CONFLICT	1170	1170	/FTID=VAR_004078.
FT	SEQUENCE	1255	137909	P -> A (IN REF. 2).
SO	SEQUENCE	1255	137909	MM; 39E9DFDA04DCE962 CRC64;

RESULT 2  
CSF2\_HUMAN STANDARD; PRT; 144 AA.

ID	CSF2_HUMAN	STANDARD;	PRT;	144	AA.
AC	P04141;				
DT	01-NOV-1986 (Rel. 03, Created)				
DT	01-NOV-1986 (Rel. 03, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)				
DE	(Colony-stimulating factor) (CSF) (Sargramostim) (Molgramostin).				
OS	CSF2 OR GMCSF.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85242684; PubMed=3925454;				
RA	Lee F., Yokota T., Otsuka T., Gemmell L., Larson N., Luh J.,				
RA	Arai K.-I., Rennick D.;				
RT	"Isolation of cDNA for a human granulocyte-macrophage				
RT	colony-stimulating factor by functional expression in mammalian				
RT	cells."				
RL	Proc. Natl. Acad. Sci. U.S.A. 82:4360-4364(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=86205844; PubMed=3486413;				
XX					



RA Kaushansky K., O'Hara P.J., Berkner K., Segal G.M., Hagen F.S.,  
 RA Adamson J.W.;  
 RT "Genomic cloning, characterization, and multilineage growth-promoting  
 RT activity of human granulocyte-macrophage colony-stimulating factor.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3101-3105(1986).  
 RL [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85298329; PubMed=3898082;  
 RA Cantrell M.A., Anderson D., Cerretti D.P., Price V., McKereghan K.,  
 RA Tushinski R.J., Mochizuki D.Y., Larsen A., Grabstein K., Cosman D.;  
 RT "Cloning, sequence, and expression of a human granulocyte/macrophage  
 RT colony-stimulating factor.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6250-6254(1985).  
 RL [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85218749; PubMed=3923623;  
 RA Wong G.G., Wittek J.S., Temple P.A., Wilkens K.M., Leary A.C.,  
 RA Luxenberg D.P., Jones S.S., Brown E.L., Kay R.M., Orr E.C.,  
 RA Shoemaker C., Golde D.W., Kaufman R.J., Hewick R.M., Wang E.A.,  
 RA Clark S.C.;  
 RT "Human GM-CSF: molecular cloning of the complementary DNA and  
 RT purification of the natural and recombinant proteins.";   
 RL Science 228:810-815(1985).  
 RL [15]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86030234; PubMed=3876930;  
 RA Miyake S., Otsuka T., Yokota T., Lee F., Arai K.-I.;  
 RT "Structure of the chromosomal gene for granulocyte-macrophage colony  
 RT stimulating factor: comparison of the mouse and human genes.";   
 RL EMBO J. 4:2561-2568(1985).  
 RL [16]  
 RP SEQUENCE FROM N.A.  
 RA Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,  
 RA Davis C.A., Kadner K., Miguel T., Pitluck S., Pollard M., Rojeski H.,  
 RA Subramanian S., Martin C.H.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RL [17]  
 RP SEQUENCE FROM N.A., AND VARIANTS ILE-115 AND THR-117.  
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
 RA Nickerson D.A.;  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RL [18]  
 RP CARBOHYDRATE-LINKAGE SITES.  
 RX MEDLINE=92144609; PubMed=1737041;  
 RA Kaushansky K., Lopez J.A., Brown C.B.;  
 RT "Role of carbohydrate modification in the production and secretion of  
 RT human granulocyte macrophage colony-stimulating factor in genetically  
 RT engineered and normal mesenchymal cells.";   
 RL Biochemistry 31:1881-1886(1992).  
 RL [19]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE=92108420; PubMed=1837174;  
 RA Diedrichs K., Boone T., Karpus P.A.;  
 RT "Novel fold and putative receptor binding site of  
 RT granulocyte-macrophage colony-stimulating factor.";   
 RL Science 254:1779-1782(1991).  
 RL [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=92235844; PubMed=1569568;  
 RA Walter M.R., Cook W.J., Ealick S.E., Nagabhushan T.L., Trotta P.P.,  
 RA Bugg C.E.;  
 RT "Three-dimensional structure of recombinant human granulocyte-  
 RT macrophage colony-stimulating factor.";   
 RL J. Mol. Biol. 224:1075-1085(1992).  
 CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION  
 CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING  
 CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- POLYMORPHISM: VARIANT ILE-117 MAY BE A RISK FACTOR FOR APOIC  
 CC ASPHMA.  
 CC -1- PHARMACEUTICAL: Available under the names Leukine (Immunex) and  
 CC Leucomax (Novartis). Used in myeloid reconstitution following bone  
 CC marrow transplant, bone marrow transplant engraftment failure or

CC delay, mobilization and following transplantation of autologous  
 CC peripheral blood progenitor cells, and following induction  
 CC chemotherapy in older adults with acute myelogenous leukemia.  
 CC -1- DATABASE: NAME-Leukine; NOTE-Clinical information on Leukine;  
 CC WWW="http://www.immunex.com/patient/pa02el.html".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M13207; AA98768.1; -  
 DR EMBL: M11734; AA52122.1; -  
 DR EMBL: M11220; AA52578.1; -  
 DR EMBL: X03021; AA26822.1; -  
 DR EMBL: M10663; AA52121.1; -  
 DR EMBL: AC004511; AAC08707.1; -  
 DR EMBL: AF373868; AA51563.1; -  
 DR PIR: A01853; F0HUGM.  
 DR PIR: C24636; C24636.  
 DR PIR: A25169; A25169.  
 DR PDB: 1CSG; 31-JAN-94.  
 DR PDB: 2GMF; 08-NOV-96.  
 DR GeneW; HGNC:2434; CSF2.  
 DR MIM: 138960; -  
 DR InterPro; IPR000773; GM-CSF.  
 DR Pfam; PF01109; GM-CSF; 1.  
 DR PRINTS; PR00693; GMCSFACTOR.  
 DR ProDom; PD007349; GM-CSF; 1.  
 DR SMART; SM00040; CSF2; 1.  
 DR PROSITE; PS00702; GM-CSF; 1.  
 KW Cytokine; Growth factor; Glycoprotein; Signal; 3D-structure;  
 KW Polymorphism; Pharmaceutical.  
 FT SIGNAL 1 17  
 FT CHAIN 18 144  
 FT  
 FT DISULFID 71 113  
 FT FT 105 138  
 FT CARBOHYD 22 22  
 FT CARBOHYD 24 24  
 FT CARBOHYD 26 26  
 FT CARBOHYD 27 27  
 FT CARBOHYD 44 44  
 FT CARBOHYD 54 54  
 FT VARIANT 115 115  
 FT  
 FT VARIANT 117 117  
 FT FT 117  
 FT TURN 25 27  
 FT HELIX 30 44  
 FT TURN 45 45  
 FT HELIX 50 54  
 FT STRAND 56 60  
 FT HELIX 72 81  
 FT TURN 82 82  
 FT HELIX 85 103  
 FT TURN 104 104  
 FT STRAND 115 119  
 FT HELIX 120 131  
 FT TURN 132 133  
 FT TURN 132 133  
 SQ SEQUENCE 144 AA; 16295 MW; 75D1E50506BCA7A8 CRC64;

Query Match 18.4%; Score 127; DB 1; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 4,2e-115;  
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 549 AARSPSPSTQWENHNAQEARLLNLSRDAENKNEVEYISEMFDQEPCTCOTRLE 608  
 DB 18 AARSPSPSTQWENHNAQEARLLNLSRDAENKNEVEYISEMFDQEPCTCOTRLE 77

Oy 609 LYKQSLRGLTKLKGPLTMASHYKOHCPMPETSCATQITFESEKENTKFLLVIPED 668  
 Db 78 LYKQSLRGLTKLKGPLTMASHYKOHCPMPETSCATQITFESEKENTKFLLVIPED 137  
 Oy 669 CMEPVQ 675  
 Db 138 CMEPVQ 144  
 RESULT 3  
 ERB2\_MESAU STANDARD: PRT: 1254 AA.  
 ID ERB2\_MESAU STANDARD: PRT: 1254 AA.  
 AC 060533:  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).  
 GN ERB2 OR NEU.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Nerve;  
 RX MEDLINE=94193007; PubMed=7908275;  
 RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,  
 RA Yamazaki Y., Ishikawa T.;  
 RT "Cloning and activation of the Syrian hamster neu proto-oncogene";  
 RL Gene 140:251-255(1994).  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,  
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
 CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
 CC (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; D16295; BAA03801.1; -  
 DR HSSP; P11362; IFGK.  
 DR InterPro: IPR000494; EGFR\_L.domain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; Pkinase.1.  
 DR Pfam: PF00757; Furin-like.1.  
 DR Pfam: PF01030; Recept\_L.domain.2.  
 DR Pfam: PF02757; YLP.2.  
 DR ProDom: PD000001; Euk\_Pkinase.1.  
 DR SMART; SM00261; FY; 3.  
 DR SMART; SM00219; TYRKC.1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR.1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM.1.  
 DR Transmembrane; Glycoprotein; MultiGene family; Receptor; Signal;  
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Proto-oncogene; Disease mutation.

FT SIGNAL 1 21  
 FT CHAIN 22 1254  
 FT DOMAIN 22 652  
 FT TRANSMEM 653 675  
 FT DOMAIN 676 1254  
 FT DOMAIN 158 368  
 FT DOMAIN 472 644  
 FT DOMAIN 720 987  
 FT NP\_BIND 726 734  
 FT BINDING 753 753  
 FT ACT\_SITE 845 845  
 FT DISULFID 195 204  
 FT DISULFID 199 212  
 FT DISULFID 236 244  
 FT DISULFID 240 252  
 FT DISULFID 255 264  
 FT DISULFID 268 295  
 FT DISULFID 299 311  
 FT DISULFID 315 331  
 FT DISULFID 334 338  
 FT DISULFID 511 520  
 FT DISULFID 515 528  
 FT DISULFID 531 540  
 FT DISULFID 544 560  
 FT DISULFID 563 576  
 FT DISULFID 567 584  
 FT DISULFID 587 596  
 FT DISULFID 600 623  
 FT DISULFID 626 634  
 FT DISULFID 630 642  
 FT MOD\_RES 1139 1139  
 FT MOD\_RES 1247 1247  
 FT CARBOHYD 68 68  
 FT CARBOHYD 125 125  
 FT CARBOHYD 187 187  
 FT CARBOHYD 259 259  
 FT CARBOHYD 530 530  
 FT CARBOHYD 571 571  
 FT CARBOHYD 629 629  
 FT CARBOHYD 658 658  
 FT VARIANT 659 659  
 FT VARIANT 659 659  
 SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F1B1 CRC64;  
 Query Match 9.78; Score 67; DB 1; Length 1254;  
 Best Local Similarity 100.0%; Pred. No. 2,8e-56;  
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 250 LPTQCHBQCAAGCTGPKHSDCLCLPHNHSIGIELHCPALVTYNTDFESMPNDEGRYT 309  
 Db 231 LPTQCHBQCAAGCTGPKHSDCLCLPHNHSIGIELHCPALVTYNTDFESMPNDEGRYT 290  
 Oy 310 FGASCVT 316  
 Db 291 FGASCVT 297  
 RESULT 4  
 ERB2\_RAT STANDARD: PRT: 1257 AA.  
 ID ERB2\_RAT STANDARD: PRT: 1257 AA.  
 AC P06494;  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor  
 DE receptor-related protein).  
 GN ERBB2 OR NEU.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.

CC TISSUE-Neuroblastoma;  
 CC MEDLINE-86118662; PubMed-3945311;  
 RA Bergmann C.I., Hung M.-C., Weinberg R.A.;  
 RT "The neu oncogene encodes an epidermal growth factor receptor-related  
 RT protein.";  
 RL Nature 319:226-230(1986).  
 RN [2]  
 RN SEQUENCE OF 852-905 FROM N.A.  
 RC TISSUE-Sciatic nerve;  
 RX MEDLINE-91222560; PubMed-2025425;  
 RA Lal C., Lemke G.;  
 RT "An extended family of protein-tyrosine kinase genes differentially  
 RT expressed in the vertebrate nervous system.";  
 RL Neuron 6:691-704(1991).  
 RN [3]  
 RN STRUCTURE BY NMR OF 650-668;  
 RX MEDLINE-92155181; PubMed-1346763;  
 RA Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,  
 RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;  
 RT "Three dimensional structure of the transmembrane region of the proto-  
 RT oncogenic and oncogenic forms of the neu protein.";  
 RL EMBO J. 11:43-48(1992).  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,  
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
 CC ALPHA AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER EBBB RECEPTORS.  
 CC THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X03362; CAA27059.1; ALT\_INIT.  
 DR PIR: A24562; TVRTNU.  
 DR HSSP: P11362; 1FGK.  
 DR InterPro: IPR000494; EGFR\_L.domain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L.domain; 2.  
 DR Pfam: PF02737; YLP; 2.  
 DR ProDom: PD000001; Euk\_Pkinase; 1.  
 DR SMART: SMO0261; FY; 3.  
 DR SMART: SMO0219; TyrKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Proto-oncogene; Disease mutation.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 1257 RECEPTOR PROTEIN-TYROSINE KINASE EBBB-2.  
 FT DOMAIN 22 654 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 655 677 POTENTIAL.  
 FT DOMAIN 678 1257 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 159 369 CYS-RICH.  
 FT DOMAIN 473 646 CYS-RICH.  
 FT DOMAIN 722 989 PROTEIN KINASE.  
 FT NP\_BIND 728 736 ATP (BY SIMILARITY).

FT BINDING 755 755 ATP (BY SIMILARITY).  
 FT ACN\_SITE 847 847 BY SIMILARITY.  
 FT DISULFID 196 205 BY SIMILARITY.  
 FT DISULFID 200 213 BY SIMILARITY.  
 FT DISULFID 221 228 BY SIMILARITY.  
 FT DISULFID 225 236 BY SIMILARITY.  
 FT DISULFID 237 245 BY SIMILARITY.  
 FT DISULFID 241 253 BY SIMILARITY.  
 FT DISULFID 256 265 BY SIMILARITY.  
 FT DISULFID 269 296 BY SIMILARITY.  
 FT DISULFID 300 312 BY SIMILARITY.  
 FT DISULFID 316 332 BY SIMILARITY.  
 FT DISULFID 335 339 BY SIMILARITY.  
 FT DISULFID 513 522 BY SIMILARITY.  
 FT DISULFID 517 530 BY SIMILARITY.  
 FT DISULFID 533 542 BY SIMILARITY.  
 FT DISULFID 546 562 BY SIMILARITY.  
 FT DISULFID 565 578 BY SIMILARITY.  
 FT DISULFID 569 586 BY SIMILARITY.  
 FT DISULFID 589 598 BY SIMILARITY.  
 FT DISULFID 602 625 BY SIMILARITY.  
 FT DISULFID 628 636 BY SIMILARITY.  
 FT DISULFID 632 644 BY SIMILARITY.  
 FT MOD\_RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).  
 SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;  
 Query Match 7.54; Score 52; DB 1; Length 1257;  
 Best Local Similarity 100.0%; Pred. No. 9e-42;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 250 LPTDCHBOCAGCTGPHKSDCLAHFNSGICGLCPALVTYNTDFESM 301  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 232 LPTDCHBOCAGCTGPHKSDCLAHFNSGICGLCPALVTYNTDFESM 283  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 5  
 PPAP\_HUMAN STANDARD; PRT; 386 AA.  
 ID PPAP\_HUMAN  
 AC P15309;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Prostatic acid phosphatase precursor (EC 3.1.3.2).  
 GN ACP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92272747; PubMed-1375464;  
 RA Sharief F.S., Li S.S.-L.;  
 RT "Structure of human prostatic acid phosphatase gene.";  
 RL Biochem. Biophys. Res. Commun. 184:1468-1476(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, DISULFIDE BONDS, AND ACTIVE  
 RP SITE.  
 RX MEDLINE-91115848; PubMed-1989985;  
 RA van Etten R.L., Davidson R., Stevis P.E., MacArthur H., Moore D.L.;  
 RT "Covalent structure, disulfide bonding, and identification of  
 RT reactive surface and active site residues of human prostatic acid  
 RT phosphatase.";  
 RL J. Biol. Chem. 266:2313-2319(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RA MEDLINE=89228054; PubMed=2712834.  
RX Sharief F.S., Lee H., Leuderman M.M., Lundwall A., Deaven L.L.,  
Lee C.-L., Li S.S.-L.;  
RT "Human prostatic acid phosphatase: cDNA cloning, gene mapping and  
protein sequence homology with lysosomal acid phosphatase.";  
RL Biochem. Biophys. Res. Commun. 160:79-86(1989).  
RN [4]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Prostate;  
RX MEDLINE=88312981; PubMed=2842184;  
RA Vilho P., Viikkuinen P., Henttu P., Roiko K., Solin T., Huhtala M.L.;  
RT "Molecular cloning and sequence analysis of cDNA encoding human  
prostatic acid phosphatase.";  
RL FEBS Lett. 236:275-281(1988).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RX MEDLINE=90370491; PubMed=2395559;  
RA Tallor P.G., Govindan M.V., Patel P.C.;  
RT "Nucleotide sequence of human prostatic acid phosphatase determined  
from a full-length cDNA clone.";  
RL Nucleic Acids Res. 18:4928-4928(1990).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95038536; PubMed=7951074;  
RA Sharief F.S., Li S.S.-L.;  
RT "Nucleotide sequence of human prostatic acid phosphatase ACPp gene,  
including seven Alu repeats.";  
RL Biochem. Mol. Biol. Int. 33:561-565(1994).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RX MEDLINE=99023966; PubMed=9804805;  
RA Lacont M.W., Handy G., Lepida L.;  
RT "Structural origins of L(+)-tartarate inhibition of human prostatic  
acid phosphatase.";  
RL J. Biol. Chem. 273:30406-30409(1998).  
CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O - an  
alcohol + phosphate.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M97589; AAA60021.1; -;  
DR EMBL; M97580; AAA60021.1; JOINED.  
DR EMBL; M97581; AAA60021.1; JOINED.  
DR EMBL; M97582; AAA60021.1; JOINED.  
DR EMBL; M97583; AAA60021.1; JOINED.  
DR EMBL; M97584; AAA60021.1; JOINED.  
DR EMBL; M97585; AAA60021.1; JOINED.  
DR EMBL; M97586; AAA60021.1; JOINED.  
DR EMBL; M97587; AAA60021.1; JOINED.  
DR EMBL; M97588; AAA60021.1; JOINED.  
DR EMBL; M34840; AAA60924.1; -;  
DR EMBL; M24902; AAA60022.1; -;  
DR EMBL; X52174; CAA36422.1; -;  
DR EMBL; X53605; CAA37673.1; -;  
DR EMBL; U07097; AAB60640.1; -;  
DR EMBL; U07083; AAB60640.1; JOINED.  
DR EMBL; U07085; AAB60640.1; JOINED.  
DR EMBL; U07086; AAB60640.1; JOINED.  
DR EMBL; U07088; AAB60640.1; JOINED.  
DR EMBL; U07091; AAB60640.1; JOINED.  
DR EMBL; U07092; AAB60640.1; JOINED.  
DR EMBL; U07093; AAB60640.1; JOINED.  
DR EMBL; U07095; AAB60640.1; JOINED.  
DR PIR; A32419; A32419.

DR PIR: S01331; S01331.  
DR PIR: S11147; S11147.  
DR PIR: JH0610; JH0610.  
DR PDB: ZHPA.16-SEP-98.  
DR Genew; HGNC:125; ACP.  
DR MIM; 171790; .  
DR InterPro: IPR000560; HisAc-phosphatase.  
DR Pfam: PF00328; acid\_phosphat\_1.  
DR PROSITE: PS00616; HIS\_ACID\_PHOSPHAT\_1; 1.  
DR PROSITE: PS00778; HIS\_ACID\_PHOSPHAT\_5; 1.  
DR KEGG: Glycolase; Glycoprotein; Signal; 3D-structure.  
FT SIGNAL 1 32  
FT CHAIN 33 386  
FT DISULFID 161 372  
FT DISULFID 215 313  
FT DISULFID 347 351  
FT ACT\_SITE 44 44  
FT ACT\_SITE 86 86  
FT ACT\_SITE 289 289  
FT CARBOHYD 94 94  
FT CARBOHYD 220 220  
FT CARBOHYD 333 333  
FT CONFLICT 15 24  
FT CONFLICT 15 24  
FT CONFLICT 15 24  
FT CONFLICT 46 46  
FT CONFLICT 66 73  
FT CONFLICT 66 73  
FT CONFLICT 95 95  
FT CONFLICT 116 116  
FT CONFLICT 139 139  
FT CONFLICT 157 157  
FT CONFLICT 212 212  
FT CONFLICT 215 215  
FT CONFLICT 294 294  
FT CONFLICT 372 372  
FT CONFLICT 383 383  
SO SEQUENCE 386 AA; 44566 MW; E81E1IDFAECADCA CRC64;  
Query Match 5.1%; Score 35; DB 1; Length 386;  
Best Local Similarity 100.0%; Pred. No. 8.9e-26;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRAAPLLARAASLSGLFLFLFFWLDNSYLAKEI 35  
DB 1 MRAAPLLARAASLSGLFLFLFFWLDNSYLAKEI 35  
RESULT 6  
CSF2\_SHEEP STANDARD: PRT: 144 AA.  
AC P28773;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)  
DE (colony-stimulating factor) (CSF).  
GN CSF2.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprine; Ovis.  
OX NCBI\_TaxID=9940;  
RX MEDLINE=92039044; PubMed=1937025;  
RT Cloning and expression of a cDNA encoding ovine  
RT granulocyte-macrophage colony-stimulating factor.";  
RL Gene1051275-279(1991).  
CC -!- FUNCTION: CITOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION  
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING  
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY

CC SIMILARITY).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>

CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).

CC -----

CC EMBL: X53561; CAA37632.1; -

CC PIR: JH0469; JH0469.

CC HSSP: P04141; 2GME.

CC InterPro: IPR000773; GM\_CSF.

CC Pfam: PF01109; GM\_CSF; 1.

CC PRINTS: PR00693; GMC\_SFCTOR.

CC PRODOM: PD007349; GM\_CSF; 1.

CC SMART: SM00040; CSF2; 1.

CC PROSITE: PS00702; GM\_CSF; 1.

CC Cytokine; Growth factor; Glycoprotein; Signal.

CC SIGNAL

CC CHAIN 1 17

CC BY SIMILARITY.

CC GRANULOCYTE-MACROPHAGE COLONY-STIMULATING

CC FACTOR.

CC DISULFID 71 113 BY SIMILARITY.

CC FT DISULFID 105 138 BY SIMILARITY.

CC FT CARBOHYD 44 44 N-LINKED (GICNAC. . .) (POTENTIAL).

CC SEQUENCE 144 AA: 16318 MW: 338580008 CRC64:

Query Match 3.2%; Score 22; DB 1; Length 144;

Best Local Similarity 100.0%; Pred. No. 1.4e-13;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 598 GEPCTQRLRLYKKGKLRGSLT 619

DB 67 GEPCTQRLRLYKKGKLRGSLT 88

RESULT 7

ERB2\_MOUSE STANDARD: PRT: 245 AA.

AC P70424; Q61525;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Receptor protein-tyrosine kinase erbB-2 (EC 2.7.1.112) (p185erbB2)

DE (NEU proto-oncogene) (C-erbB-2) (fragments).

GN ERBB2 OR NEU.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_Taxid=10090;

RN [1]

RN SEQUENCE OF 1-149 FROM N.A.

RP STRAIN-CD-1; TISSUE-Uterus;

RX MEDLINE-97200814; PubMed-9048643;

RA Lin J., Day S.K., Das S.K.;

RT "Differential expression of the erbB2 gene in the perimplantation

RT mouse uterus: potential mediator of signaling by epidermal growth

RT factor-like growth factors.";

RL Endocrinology 138:1328-1337(1997).

RN [2]

RP SEQUENCE OF 150-245 FROM N.A.

RX MEDLINE-96069911; PubMed-7589796;

RA Moscoso L.M., Chu G.C., Gautam M., Noakes P.G., Merlie J.P.,

RA Sanes J.R.;

RT "Synapse-associated expression of an acetylcholine receptor-inducing

RT protein, ARIA/heretulin, and its putative receptors, ErbB2 and ErbB3,

RT in developing mammalian muscle.";

RL Dev. Biol. 172:158-169(1995).

CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,

CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A

CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-

CC ALPHA AND AMPHIREGULIN.

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein

CC tyrosine phosphate.

CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS

CC (POTENTIAL).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN UTERINE EPITHELIAL

CC CELLS. IN THE MUSCLE, EXPRESSION LOCALIZES TO THE SYNAPTIC SITES

CC OF MUSCLE FIBERS.

CC -1- DEVELOPMENTAL STAGE: ON DAYS 1-4 OF PREGNANCY, ERBB2 IS DETECTED

CC PRIMARILY IN EPITHELIAL CELLS, THE DAY 1 UTERUS SHOWING THE

CC HIGHEST ACCUMULATION. ON DAY 5, THE EPITHELIUM AND THE

CC DECIDUALIZING STROMAL CELLS AROUND THE IMPLANTING BLASTOCYST

CC EXHIBIT ACCUMULATION OF THIS RECEPTOR. ON DAYS 6-8, THE EXPRESSION

CC PERSISTS IN THE EPITHELIUM AT BOTH THE IMPLANTATION AND

CC INTERIMPLANTATION SITES IN ADDITION TO MODEST LEVELS IN THE

CC SECONDARY DECIDUAL ZONE. ON DAYS 7 AND 8, ACCUMULATION IS ALSO

CC PROMINENT IN THE TROPHOBLASTIC GIANT CELLS.

CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE

CC RESIDUES (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>

CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).

CC -----

CC EMBL: U71126; AAB17380.1; -

CC EMBL: L47239; AAB93532.1; -

CC HSSP: P11362; 1PGR.

CC MGD: M61:95410; ErbB2.

CC InterPro: IPR000719; Euk\_pkinase.

CC InterPro: IPR004040; STY\_pkinase.

CC InterPro: IPR001245; TYR\_pkinase.

CC PRINTS: PR00109; TYRKINASE.

CC PRODOM: PD000001; Euk\_pkinase; 1.

CC SMART: SM00221; STYK; 1.

CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; PARTIAL.

CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.

CC PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.

CC Transmembrane; Glycoprotein; Multigene family; Receptor; Transferase;

CC Tyrosine-protein kinase; ATP-binding; Phosphorylation.

CC NON\_TER 1 1

CC DOMAIN 1 >149

CC ACT\_SITE 61 61

CC NON\_CONS 149 150

CC NON\_TER 245 245

CC SEQUENCE 245 AA: 26927 MW: 576310363DFEFIC CRC64:

Query Match 2.9%; Score 20; DB 1; Length 245;

Best Local Similarity 100.0%; Pred. No. 1.9e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 363 PRSPAPSGAGSDVFDGDL 382

DB 210 PRSPAPSGAGSDVFDGDL 229

RESULT 8

CSF2\_CAVPO STANDARD: PRT: 140 AA.

ID CSF2\_CAVPO

AC Q60481;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)

DE (Colony-stimulating factor) (CSF) (Fragment).

GN CSF2.

OS Cavia porcellus (Guinea pig).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

CC Mammalia; Euthera; Rodentia; Hystriognathi; Cavidae; Cavia.
CX NCBI_TaxID-10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Hartley; Tissue-Lung;
RA Yuan H.T., Kelly F.J., Bingle C.D.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U46779; AAA87592.1; -
DR HSSP; P04141; ZGMF.
DR InterPro: IPR000773; GM_CSF.
DR Pfam: PF01109; GM_CSF; 1.
DR ProDom: PD007349; GM_CSF; 1.
DR SMART; SM00040; CSF2; 1.
DR PROSITE; PS00702; GM_CSF; 1.
DR CYTOKINE; Growth factor; glycoprotein; Signal.
KW CYTOKINE; Growth factor; glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 >140
FT FT GRANULOCYTE-MACROPHAGE COLONY-STIMULATING
FT FT FACTOR.
FT FT BY SIMILARITY.
FT FT DISULFID 72 114
FT FT 106 139
FT FT CARBOHYD 45 45
FT FT CARBOHYD 55 55
FT FT CARBOHYD 87 87
FT FT NON_PER 140 140
FT FT SEQUENCE 140 AA; 15869 MW; FF72E7A9491C831C CRC64;

Query Match 1.9%; Score 13; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 YKHCPCPPPTSC 644
Db 102 YKHCPCPPPTSC 114

RESULT 9
CSF2_CANFA STANDARD; PRT; 144 AA.
ID CSF2_CANFA
AC P48749;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
DE (Colony-stimulating factor) (CSF).
GN CSF2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID-9615;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-91329842; Pubmed-1868252;
RA Nash R.A., Schenning F., Appelbaum F., Hammond W.P., Boone T.,
RA Morris C.F., Slichter S.J., Stord R.;
RT "Molecular cloning and in vivo evaluation of canine granulocyte-
RT macrophage colony-stimulating factor.";
RL Blood 78:930-937(1991).
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING

```

```

CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S49738; AAB19466.1; -
DR HSSP; P04141; ZGMF.
DR InterPro: IPR000773; GM_CSF.
DR Pfam; PF01109; GM_CSF; 1.
DR PRINTS; PR00693; GMCSFACOR.
DR ProDom: PD007349; GM_CSF; 1.
DR SMART; SM00040; CSF2; 1.
DR PROSITE; PS00702; GM_CSF; 1.
DR CYTOKINE; Growth factor; glycoprotein; Signal.
KW CYTOKINE; Growth factor; glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 144
FT FT GRANULOCYTE-MACROPHAGE COLONY-STIMULATING
FT FT FACTOR.
FT FT BY SIMILARITY.
FT FT DISULFID 71 113
FT FT 105 138
FT FT CARBOHYD 44 44
FT FT SEQUENCE 144 AA; 16137 MW; 6323807E1FC6343 CRC64;

Query Match 1.6%; Score 11; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 631 HYKHCPTPE 641
Db 100 HYKHCPTPE 110

RESULT 10
CSF2_CEREL STANDARD; PRT; 144 AA.
ID CSF2_CEREL
AC P51748;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
DE (Colony-stimulating factor) (CSF).
GN CSF2.
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID-9860;
RN [1]
RP SEQUENCE FROM N.A.
RA Lockhart E.A.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14392; AAA21439.1; -
DR HSSP; P04141; ZGMF.

```

```
DR InterPro; IPR000773; GM_CSF.
DR Pfam; PF01109; GM_CSF; 1.
DR PRINTS; PR00693; GMCSFACOR.
DR PRODOM; PD007349; GM_CSF; 1.
DR SMART; SM00040; CSF2; 1.
DR PROSITE; PS00702; GM_CSF; 1.
DR Cyclicline; Growth factor; Glycoprotein; Signal.
KW CYCLICLINE; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 144
FT DISULFID 71 113
FT DISULFID 105 138
FT CARBOHYD 44 44
FT CARBOHYD 54 54
SQ SEQUENCE 144 AA; 16283 MM; 1F5FE5FD03C94394 CRC64;

Query Match
Best Local Similarity 1.68; Score 11; DB 1; Length 144;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 609 LYKGLRGSLSLT 619
DB 78 LYKGLRGSLSLT 88
```

```
RESULT 11
CSF2_PIG STANDARD: PRT: 144 AA.
AC 029118; Q29046;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
DE (Colony-stimulating factor) (CSF).
GN CSF2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=96823;
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Foss D.L., Murtaugh M.P.;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RA Gloster S.E., Sandeman R.M., Strom A.D.G.;
RL Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RA MEDLINE=96167041; PubMed=8595928;
RA Inumaru S., Takamatsu H.;
RT "cDNA cloning of porcine granulocyte-macrophage colony-stimulating
RT factor.";
RL Immunol. Cell Biol. 73:474-476(1995).
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -1- SUBUNIT: MONOMER.
CC -1- SUBUNIT: MONOMER.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U61139; AAB03867.1; -
CC EMBL; U67318; AAB49939.1; -
CC EMBL; U67175; AAB06854.1; -
CC EMBL; D21074; BAA04649.1; -
```

```
DR HSSP; P04141; 2GME.
DR InterPro; IPR000773; GM_CSF.
DR Pfam; PF01109; GM_CSF; 1.
DR PRINTS; PR00693; GMCSFACOR.
DR PRODOM; PD007349; GM_CSF; 1.
DR SMART; SM00040; CSF2; 1.
DR PROSITE; PS00702; GM_CSF; 1.
DR Cyclicline; Growth factor; Glycoprotein; Signal.
KW CYCLICLINE; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 144
FT DISULFID 71 113
FT DISULFID 105 138
FT CARBOHYD 44 44
FT CARBOHYD 47 47
FT CARBOHYD 54 54
FT CARBOHYD 54 54
FT CONFLICT 59 59
FT CONFLICT 140 140
FT CONFLICT 142 143
SQ SEQUENCE 144 AA; 16254 MM; 793DACB62CF736D0 CRC64;

Query Match
Best Local Similarity 1.68; Score 11; DB 1; Length 144;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 609 LYKGLRGSLSLT 619
DB 78 LYKGLRGSLSLT 88
```

```
RESULT 12
CSF2_BOVIN STANDARD: PRT: 143 AA.
AC P11032;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
DE (Colony-stimulating factor) (CSF).
GN CSF2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RC SEQUENCE FROM N.A.
RA MEDLINE=89096971; PubMed=3062386;
RA Maliszewski C.R., Schoenborn M.A., Cerretti D.P., Wignall J.M.,
RA Picha K.S., Cosman D., Tushinski R.J., Gillis S., Baker P.E.;
RT "Bovine GM-CSF: molecular cloning and biological activity of the
RT recombinant protein.";
RL Mol. Immunol. 25:843-850(1988).
RN [2]
RC SEQUENCE FROM N.A.
RA Leong S.R., Flagg G.M., Lawman M.J.P., Gray P.W.;
RT "Cloning and expression of the cDNA for bovine granulocyte-macrophage
RT colony-stimulating factor.";
RL Vet. Immunol. Immunopathol. 21:261-278(1989).
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

CC EMBL: U22385; AAA6075.1; -  
 DR PIR: J10037; F0B0GM.  
 DR HSSP: P04141; 2GMF.  
 DR InterPro: IPR000773; GM\_CSF.  
 DR Pfam: PF01109; GM\_CSF; 1.  
 DR PRINTS: PR00693; GMCSFACTOR.  
 DR PRODOM: PD007349; GM\_CSF; 1.  
 DR SMART: SM00040; CSF2; 1.  
 DR PROSITE: PS00702; GM\_CSF; 1.  
 KM Cytokine; growth factor; glycoprotein; signal.  
 FT SIGNAL 1 17  
 FT CHAIN 18 143  
 FT FT 18 143 GRANULOCYTE-MACROPHAGE COLONY-STIMULATING  
 FT DISULFID 70 112 FACTOR.  
 FT DISULFID 104 137 BY SIMILARITY.  
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 143 AA; 16157 MW; 4A24E26A870A51EC CRC64;

Query Match 1.48; Score 10; DB 1; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 0.057;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 QEPICLQTRL 607  
 DB 66 QEPICLQTRL 75

RESULT 13  
 ZIC2\_MOUSE STANDARD; PRT; 530 AA.  
 AC 062520;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Zinc finger protein ZIC2 (zinc finger protein of the cerebellum 2).  
 GN ZIC2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=96132843; PubMed=8557628;  
 RA Aruga J., Nagai T., Tokuyama T., Hayashizaki Y., Okazaki Y.,  
 RA Chapman V.M., Mikoshiba K.,  
 RT "The mouse zic gene family. Homologues of the Drosophila pair-rule  
 RT gene odd-paired.";  
 RL J. Biol. Chem. 271:1043-1047(1996).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: CNS. A HIGH LEVEL EXPRESSION IS SEEN IN THE  
 CC CEREBELLUM.  
 CC -1- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER  
 CC PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: D70848; BAAL115.1; -  
 DR HSSP: P08047; 1SP2.  
 DR TRANSFAC: T04670; -  
 DR MGD: MGI:106679; ZIC2.  
 DR InterPro: IPR000822; znf\_C2H2.  
 DR Pfam: PF00096; znf\_C2H2; 5.  
 DR PRINTS: PR00046; ZINC2FINGER.  
 DR PRODOM: PD000003; znf\_C2H2; 1.

DR SMART: SM00355; znf\_C2H2; 4.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 3.  
 DR PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2; 4.  
 KM Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.  
 FT DOMAIN 20 23 POLY-HIS.  
 FT DOMAIN 25 33 POLY-ALA.  
 FT DOMAIN 89 97 POLY-ALA.  
 FT DOMAIN 227 231 POLY-ALA.  
 FT DOMAIN 232 239 POLY-HIS.  
 FT DOMAIN 300 415 ZINC FINGERS.  
 FT ZN\_FING 300 327 ZINC FINGERS.  
 FT ZN\_FING 333 357 C2H2-TYPE.  
 FT ZN\_FING 363 387 C2H2-TYPE.  
 FT ZN\_FING 393 415 C2H2-TYPE.  
 FT ZN\_FING 456 470 POLY-ALA.  
 FT DOMAIN 501 512 POLY-ALA.  
 SQ SEQUENCE 530 AA; 55492 MW; 0065BD75B52E7DD2 CRC64;

Query Match 1.38; Score 9; DB 1; Length 530;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 682 AAAHHHHH 690  
 DB 229 AAAHHHHH 237

RESULT 14  
 ZIC2\_HUMAN STANDARD; PRT; 532 AA.  
 AC 095409; O9H309.  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Zinc finger protein ZIC2 (zinc finger protein of the cerebellum 2).  
 GN ZIC2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND VARIANT HPES POLY-ALA INSERTION.  
 RX MEDLINE=98442655; PubMed=9771712;  
 RA Brown S.A., Warburton D., Brown L.Y., Yu C.Y., Roeder E.R.,  
 RA Stengel-Rutkowski S., Hennekam R.C., Muenke M.,  
 RT "Holoencephaly due to mutations in ZIC2, a homologue of Drosophila  
 RT odd-paired.";  
 RL Nat. Genet. 20:180-183(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20556339; PubMed=10984499;  
 RA Yang Y., Hwang C.K., Junn E., Lee G., Mouradian M.M.,  
 RT "ZIC2 and Sp3 repress Sp1-induced activation of the human D1A dopamine  
 RT receptor gene.";  
 RL J. Biol. Chem. 275:38863-38869(2000).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DISEASE: DEFECTS IN ZIC2 ARE A CAUSE OF HOLOPROSEPHALY TYPE 5  
 CC (HPES). HPES IS A STRUCTURAL ANOMALY OF THE BRAIN.  
 CC -1- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER  
 CC PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF104902; AAC96325.1; -  
 DR EMBL: AF193855; AAC28409.1; -  
 DR HSSP: P08047; 1SP2.



DR TRANSFAC; T04237; .  
 DR Genew; HGNC:12873; ZIC2.  
 DR MIM: 603073; .  
 DR InterPro; IPR000822; ZnF\_C2H2.  
 DR Pfam; PF00096; ZF-C2H2; 5.  
 DR ProDom; PD000003; ZnF\_C2H2; 1.  
 DR SMART; SM00355; ZnF\_C2H2; 4.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 3.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 4.  
 DR Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;  
 KW Disease mutation; Holoprosencephaly.  
 KM  
 FT DOMAIN 20  
 FT DOMAIN 23  
 FT DOMAIN 25  
 FT DOMAIN 89  
 FT DOMAIN 97  
 FT DOMAIN 226  
 FT DOMAIN 230  
 FT DOMAIN 231  
 FT DOMAIN 239  
 FT ZN\_FING 300  
 FT ZN\_FING 300  
 FT ZN\_FING 327  
 FT ZN\_FING 333  
 FT ZN\_FING 357  
 FT ZN\_FING 363  
 FT ZN\_FING 387  
 FT ZN\_FING 393  
 FT ZN\_FING 415  
 FT ZN\_FING 456  
 FT ZN\_FING 470  
 FT DOMAIN 490  
 FT DOMAIN 508  
 FT VARIANT 470  
 FT CONFLICT 124  
 FT SEQUENCE 532 AA; 128  
 Query Match  
 Best Local Similarity 1.3%; Score 9; DB 1; Length 532;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 682 AAAAAHHHHH 690  
 DB 228 AAAAAHHHHH 236

RESULT 15  
 EGFRC\_CHICK  
 ID EGFRC\_CHICK STANDARD; PRT; 703 AA.  
 AC P13387;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)  
 DE (Fragment).  
 GN EGFRC.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE:86261272; PubMed:3260329;  
 RA Lax I., Johnson A., Howk R., Sap J., Bellot F., Winkler M.,  
 RA Ullrich A., Vennstrom B., Schlessinger J., Givol D.;  
 RA "Chicken epidermal growth factor (EGF) receptor: cDNA cloning,  
 RT expression in mouse cells, and differential binding of EGF and  
 RT transforming growth factor alpha".  
 RL Mol. Cell. Biol. 8:1970-1978(1988).  
 CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,  
 CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND  
 CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to  
 CC dimerization, internalization of the EGF-receptor complex,  
 CC induction of the tyrosine kinase activity, stimulation of cell DNA  
 CC synthesis, and cell proliferation.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M20386; AAAA8760.1; .  
 DR InterPro; IPR000494; EGFRC\_L\_domain.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002174; Furin-like.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR SMART; SM00261; FU; 4.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; PARTIAL.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; PARTIAL.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; PARTIAL.  
 KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;  
 KW Tyrosine-protein kinase; ATP-binding; phosphorylation.  
 FT SIGNAL 1  
 FT CHAIN 1  
 FT DOMAIN 31  
 FT DOMAIN 31  
 FT TRANSMEM 655  
 FT DOMAIN 668  
 FT DISULFID 197  
 FT DISULFID 201  
 FT DISULFID 222  
 FT DISULFID 222  
 FT DISULFID 226  
 FT DISULFID 239  
 FT DISULFID 243  
 FT DISULFID 258  
 FT DISULFID 271  
 FT DISULFID 271  
 FT DISULFID 302  
 FT DISULFID 302  
 FT DISULFID 318  
 FT DISULFID 336  
 FT DISULFID 340  
 FT DISULFID 513  
 FT DISULFID 517  
 FT DISULFID 533  
 FT DISULFID 542  
 FT DISULFID 546  
 FT DISULFID 565  
 FT DISULFID 569  
 FT DISULFID 592  
 FT DISULFID 605  
 FT DISULFID 630  
 FT DISULFID 634  
 FT CARBOHYD 134  
 FT CARBOHYD 134  
 FT CARBOHYD 190  
 FT CARBOHYD 200  
 FT CARBOHYD 200  
 FT CARBOHYD 359  
 FT CARBOHYD 368  
 FT CARBOHYD 368  
 FT CARBOHYD 420  
 FT CARBOHYD 420  
 FT CARBOHYD 573  
 FT CARBOHYD 578  
 FT CARBOHYD 613  
 FT CARBOHYD 633  
 FT CARBOHYD 648  
 FT NON\_TER 703  
 FT SEQUENCE 703 AA; 77427 MW; AAF2DE11B735690 CRC64;  
 Query Match  
 Best Local Similarity 1.3%; Score 9; DB 1; Length 703;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266  
 DB 242 QCAAGCTGP 250

Search completed: April 28, 2003, 13:44:58  
 Job time : 17 secs

**THIS PAGE BLANK (USPTO)**



```

QY 98 EVQGVLIANNOVROVPLQRLIRVGTQLEFEDNALAVLNDGDPINNTPTVTGASPGCLR 157
DB 79 EVQGVLIANNOVROVPLQRLIRVGTQLEFEDNALAVLNDGDPINNTPTVTGASPGCLR 138
QY 158 ELQURSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPM 217
DB 139 ELQURSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPM 198
QY 218 CKGRSGESEDOSLITRVNCGGCAKCGPLPTDCCHQCAAGCPHSDCLACHF 277
DB 199 CKGRSGESEDOSLITRVNCGGCAKCGPLPTDCCHQCAAGCPHSDCLACHF 258
QY 278 NHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNYLSTDVGS 329
DB 259 NHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNYLSTDVGS 310

RESULT 2
Q14256 PRELIMINARY: PRT: 165 AA.
ID Q14256:
AC Q14256:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE C-erb B2/neu protein (Fragment).
GN C-ERB B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6070181; Pubmed=2999974;
RA Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,
RA Seeburg P.H., Liberman T.A., Schlessinger J., Franke U.,
RA Levison A., Ullrich A.;
RT Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.;
RL Science 230:1132-1139(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94000386; Pubmed=8104414;
RA Sarkar F.H., Ball D.E., Li Y.W., Cissman J.D.;
RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERB2)
RT gene.";
RL DNA Cell Biol. 12:611-615(1993).
DR EMBL: M95667; AAC37531.1;
FT NON_TER 1
SQ SEQUENCE 165 AA; 17327 MW; A0C113BA308BF46B CRC64;

Query Match 23.9%; Score 165; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 2, 2e-159;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 373 AGSDVFPFGDLGMAKAGLQSLPHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV 432
DB 1 AGSDVFPFGDLGMAKAGLQSLPHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV 60
QY 433 NOPDVRRQPPSPREGPLPAARPAAGATLERAKTISPGKNGVAKDVFAGGAVENPEYLTQP 492
DB 61 NOPDVRRQPPSPREGPLPAARPAAGATLERAKTISPGKNGVAKDVFAGGAVENPEYLTQP 120
QY 493 GGAAPQHPHPAPASPAEDNLTYMDQDPERGAPESTFKGTPTAEN 537
DB 121 GGAAPQHPHPAPASPAEDNLTYMDQDPERGAPESTFKGTPTAEN 165

RESULT 3
Q8WYV0 PRELIMINARY: PRT: 412 AA.
ID Q8WYV0:
AC Q8WYV0:
DT 01-MAR-2002 (TREMblrel. 20, Created)

```

```

DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 44.7 kDa protein.
GN P3659.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.O., Qin W.X., Zhao X.T.,
RA Wan D.F., Gu J.R.;
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF18349; AAL5856.1;
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR00719; Euk_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PFO0069; VLP_motif.
DR Pfam: PFO2757; VLP_2.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KM Hypothetical protein.
SQ SEQUENCE 412 AA; 44702 MW; 034397FF3F27D2BC CRC64;

Query Match 19.18%; Score 132; DB 4; Length 412;
Best Local Similarity 100.0%; Pred. No. 2, 2e-125;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 GAGGMVHRRSSSTRSGGDLTLGLPSEEARSPPLAPSEGASDVFQDLGMAKAG 389
DB 150 GAGGMVHRRSSSTRSGGDLTLGLPSEEARSPPLAPSEGASDVFQDLGMAKAG 209
QY 390 LQSLPTHDPSPLOKRYSEDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRRQPPSPREGPL 449
DB 210 LQSLPTHDPSPLOKRYSEDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRRQPPSPREGPL 269
QY 450 PAARPAAGATLER 461
DB 270 PAARPAAGATLER 281

RESULT 4
Q9BG66 PRELIMINARY: PRT: 149 AA.
ID Q9BG66:
AC Q9BG66:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE Receptor tyrosine kinase ErbB2 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
RA Tectens F., Fischer B.;
RT "ErbB genes and epidermal growth factor- (EGF-) like ligands in the
RT sub-implantation rabbit uterus and blastocyst.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF333178; AAK14371.1;
DR InterPro: IPR002174; Furin-like.
DR Pfam: PFO0757; Furin-like; 1.
DR SMART: SM00261; FU; 2.
DR Kinase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;

```

```

Query Match      9.7%; Score 67; DB 6; Length 149;
Best Local Similarity 100.0%; Pred. No. 1.3e-59;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 250 LPTDCHEQCAAGCTGPRKSDCLACLFHNSGICELCPALVTYNTDTFESMPNEGRTY 309
    |||||||
DB 74 LPTDCHEQCAAGCTGPRKSDCLACLFHNSGICELCPALVTYNTDTFESMPNEGRTY 133
    |||||||

OY 310 FGASCVT 316
    |||||||
DB 134 FGASCVT 140

RESULT 5
ID 018735 PRELIMINARY; PRT; 1259 AA.
AC 018735;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ErbB-2.
OS Carls familiaris (dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.;
RT "cDNA cloning of erbB-2 from canine mammary gland."
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB008451; BAA23127.1;
DR HSSP; P11362; IFGK.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002174; Furin-like.
DR InterPro; IPR001245; Tyr-kinase.
DR InterPro; IPR004019; YLP_motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR SMART; SM00261; FU; 3.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1259 AA; 137968 MW; E37364D49C4ACD46 CRC64;

Query Match      9.6%; Score 66; DB 6; Length 1259;
Best Local Similarity 100.0%; Pred. No. 1.1e-57;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 PTDCHEQCAAGCTGPRKSDCLACLFHNSGICELCPALVTYNTDTFESMPNEGRTY 310
    |||||||
DB 232 PTDCHEQCAAGCTGPRKSDCLACLFHNSGICELCPALVTYNTDTFESMPNEGRTY 291
    |||||||

OY 311 GASCVT 316
    |||||||
DB 292 GASCVT 297

RESULT 6
ID 09GL44 PRELIMINARY; PRT; 144 AA.
AC 09GL44;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

```

```

DE Granulocyte-macrophage colony-stimulating factor.
GN GM-CSF.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Hutchinson K.L., Villinger F., Miranda M.E., Ksiazek T.G.,
RA Peters C.J., Rollin P.E.;
RT "Multiplex analysis of cytokines in the sera of cynomolgus macaques
RT naturally infected with Ebola (Reston)."
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY007376; AAG16626.1;
DR HSSP; P04141; 2GMF.
DR InterPro; IPR000773; GM-CSF.
DR Pfam; PF01109; GM-CSF; 1.
DR PRINTS; PR00693; GMCSFACTOR.
DR ProDom; PD007349; GM-CSF; 1.
DR SMART; SM00040; CSF2; 1.
DR PROSITE; PS00702; GM-CSF; 1.
FT VARIANT 60 60 V -> I.
SQ SEQUENCE 144 AA; 16177 MW; 7D5F381DA2FC832P CRC64;

Query Match      5.2%; Score 36; DB 6; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.8e-28;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 GSLTKLKGPLTWMASHYKHCPRPTSCAQITTF 651
    |||||||
DB 85 GSLTKLKGPLTWMASHYKHCPRPTSCAQITTF 120
    |||||||

RESULT 7
ID 096KY0 PRELIMINARY; PRT; 386 AA.
AC 096KY0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Acid phosphatase, prostatic.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Prostate;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC016344; AAH16344.1;
DR InterPro; IPR000560; HsAc_phsptase.
DR Pfam; PF00328; acid_phsphat; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ SEQUENCE 386 AA; 44540 MW; FE30E10CBCECDA CRC64;

Query Match      5.1%; Score 35; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.3e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRAAPLLARRASTLSGFLFLFPLDRSVLAKEL 35
    |||||||
DB 1 MRAAPLLARRASTLSGFLFLFPLDRSVLAKEL 35
    |||||||

RESULT 8
ID 096QMO PRELIMINARY; PRT; 418 AA.
AC 096QMO;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

```

DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)  
 DE Acid phosphatase, prostate.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PROSTATE;  
 RA Strauberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC007460; AAH07460.1; -  
 DR InterPro: IPR000560; HsAc.phosphatase.  
 DR Pfam: Pf00328; acid\_phosphatase.1.  
 DR PROSITE: PS00616; HIS\_ACID\_PHOSPHAT\_1; UNKNOWN.1.  
 DR PROSITE: PS00778; HIS\_ACID\_PHOSPHAT\_2; UNKNOWN.1.  
 SO SEQUENCE 418 AA; 48308 MW; 168E10406974E4462 CRC64;

Query Match 5.1%; Score 35; DB 4; Length 418;  
 Best Local Similarity 100.0%; Pred. No. 1,4e-26;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRAAPLLARAASLSLGFLLFFLFDRLSYLAKEL 35  
 Db 1 MRAAPLLARAASLSLGFLLFFLFDRLSYLAKEL 35

RESULT 9  
 ID 08R2X1 PRELIMINARY; PRT; 367 AA.  
 AC 08R2X1;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)  
 DE Hypothetical 40.2 kDa protein.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strauberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC027080; AAH27080.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 367 AA; 40163 MW; 0BE03395F9E101B0 CRC64;

Query Match 3.9%; Score 27; DB 11; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-18;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 463 KTLSPGKNGVYKDYFAFGAVENPEYL 489  
 Db 283 KTLSPGKNGVYKDYFAFGAVENPEYL 309

RESULT 10  
 ID 09MYK4 PRELIMINARY; PRT; 144 AA.  
 AC 09MYK4;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)  
 DE Granulocyte-macrophage colony-stimulating factor (Fragment).  
 OS Ovis aries (Sheep).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprine; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91331592; PubMed=1869289;  
 RA O'Brien P.M., Rothel J.S., Seow H.F., Wood P.R.;

RT "Cloning and sequencing of the cDNA for ovine granulocyte-macrophage  
 colony stimulating factor (GM-CSF)."  
 RL Immunol. Cell Biol. 69:51-55(1991).  
 DR EMBL: X55991; CAA39463.1; -  
 DR HSSP: P04141; 2GMF  
 DR InterPro: IPR000773; GM-CSF.  
 DR Pfam: Pf01109; GM-CSF.1.  
 DR PRINTS: PR00693; GMCSFACTOR.  
 DR PRODOM: PD007349; GM-CSF.1.  
 DR SMART: SM00040; CSF2.1.  
 DR PROSITE: PS00702; GM-CSF.1.  
 FT NON\_TER  
 SO SEQUENCE 144 AA; 16290 MW; ABAAD7633B580008 CRC64;

Query Match 3.2%; Score 22; DB 6; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 QEPICLQTRLELYKQGLRSLT 619  
 Db 67 QEPICLQTRLELYKQGLRSLT 88

RESULT 11  
 ID 096QK9 PRELIMINARY; PRT; 386 AA.  
 AC 096QK9;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)  
 DE Acid phosphatase, prostate.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PROSTATE;  
 RA Strauberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC008493; AAH08493.1; -  
 DR InterPro: IPR000560; HsAc.phosphatase.  
 DR Pfam: Pf00328; acid\_phosphatase.1.  
 DR PROSITE: PS00616; HIS\_ACID\_PHOSPHAT\_1; UNKNOWN.1.  
 DR PROSITE: PS00778; HIS\_ACID\_PHOSPHAT\_2; UNKNOWN.1.  
 SQ SEQUENCE 386 AA; 44515 MW; AAD817CEC1DC1A84 CRC64;

Query Match 2.9%; Score 20; DB 4; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-11;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LGFLFLFFWLDRLSYLAKEL 35  
 Db 16 LGFLFLFFWLDRLSYLAKEL 35

RESULT 12  
 ID 08WN17 PRELIMINARY; PRT; 146 AA.  
 AC 08WN17;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)  
 DE Granulocyte-macrophage colony-stimulating factor (Fragment).  
 OS Equus caballus (Horse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vecchiarelli A., D'Amico F., Kanellos T.S., Howard C.R., Hamblin A.S.,  
 Catchpole B.;

RT "cDNA for equine GM-CSF."  
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF448481; AAL1017.1;  
 DR InterPro: IPR000773; GM-CSF.  
 DR Pfam: PF01109; GM-CSF; 1.  
 DR PRINTS: PR00693; GMCSFACTOR.  
 DR ProDom: PD007349; GM-CSF; 1.  
 DR SMART: SM00040; CSF2; 1.  
 DR PROSITE: PS00702; GM-CSF; UNKNOWN\_1.  
 FT NON-TER 146 146  
 SQ SEQUENCE 146 AA: 16594 MW: 162C19152391281E CRC64;

Query Match 2.5%; Score 17; DB 6; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 623 GPLTMASHYKHCPT 639  
 DB 92 GPLTMASHYKHCPT 108

RESULT 13  
 O95L10 PRELIMINARY: PRT; 152 AA.

AC 095L10;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Granulocyte-macrophage colony-stimulating-factor.  
 GN GM-CSF.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Manel S., Commandeur U., Steinhach F.;  
 RT "Cloning of equine granulocyte-macrophage colony-stimulating-factor  
 (eq-GM-CSF)."  
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AY040203; AAK72108.2;  
 DR InterPro: IPR000773; GM-CSF.  
 DR Pfam: PF01109; GM-CSF; 1.  
 DR ProDom: PD007349; GM-CSF; 1.  
 DR PROSITE: PS00702; GM-CSF; UNKNOWN\_1.  
 SQ SEQUENCE 152 AA: 17173 MW: 75605CC1ADE9EEF9 CRC64;

Query Match 2.5%; Score 17; DB 6; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 623 GPLTMASHYKHCPT 639  
 DB 92 GPLTMASHYKHCPT 108

RESULT 14  
 O99J91 PRELIMINARY: PRT; 138 AA.

AC 099J91;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Granulocyte-macrophage colony stimulating factor precursor.  
 OS Marmota monax (Woodchuck).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;  
 OC Marmota.  
 OX NCBI\_TaxID=9995;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wu H.-L., Chen P.-J., Lin H.-K., Lee R.-S., Lin H.-L., Chen D.-S.;  
 RT "Molecular Cloning and Expression of Woodchuck Granulocyte-Macrophage

RT Colony Stimulating Factor."  
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF255734; AAG49541.1;  
 DR EMBL: AF255735; AAG49542.1;  
 DR HSRF: P04141; 2GMF.  
 DR InterPro: IPR000773; GM-CSF.  
 DR Pfam: PF01109; GM-CSF; 1.  
 DR PRINTS: PR00693; GMCSFACTOR.  
 DR ProDom: PD007349; GM-CSF; 1.  
 DR SMART: SM00040; CSF2; 1.  
 DR PROSITE: PS00702; GM-CSF; 1.  
 FT SIGNAL.  
 FT SIGNAL.  
 FT CHAIN 20 138  
 FT CHAIN 20 138  
 SQ SEQUENCE 138 AA: 15297 MW: A6B9E04035A806FD CRC64;

Query Match 1.3%; Score 9; DB 11; Length 138;  
 Best Local Similarity 100.0%; Pred. No. 1.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 599 EPTCLOTRL 607  
 DB 68 EPTCLOTRL 76

RESULT 15  
 O8XU41 PRELIMINARY: PRT; 253 AA.

AC O8XU41;  
 DT 01-MAR-2002 (TREMblrel. 20, Created)  
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Hypothetical transmembrane protein RSC3353.  
 GN RSC3353 OR RS02628.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GM11000;  
 RX MEDLINE-21681879; PubMed-11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Ariat M., Billault A., Brotlier P., Camus J.C., Cartolico L.,  
 RA Chandlier M., Choise N., Claudel-Renard C., Cunac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Siglier P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Welzenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646074; CAD17141.1;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 253 AA: 2668 MW: 95DBE888F9F0B58 CRC64;

Query Match 1.3%; Score 9; DB 16; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 2.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LLLARRASL 14  
 DB 68 LLLARRASL 76

Search completed: April 28, 2003, 13:45:48  
 Job time : 47 secs

**THIS PAGE BLANK (USPTO)**



QY 98 EVGGYVLIAHNQVRQVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGCLR 15

Db 79 EVGGVLIANHNOVQVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 138  
Qy 158 ELQLSRLEILKGVLIQIORNPOLCYODITLWKDIFHKNNQALTLIDITNRSRACHPCSPM 217  
Db 139 ELQLSRLEILKGVLIQIORNPOLCYODITLWKDIFHKNNQALTLIDITNRSRACHPCSPM 198  
Qy 218 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 277  
Db 199 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 258  
Qy 278 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNLTSDVGS 329  
Db 259 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNLTSDVGS 310

RESULT 2  
US-09-146-283-4  
Sequence 4, Application US/09146283  
Patent No. 5976546

GENERAL INFORMATION:  
APPLICANT: Laus, Reiner  
APPLICANT: Ruegg, Curtis L.  
APPLICANT: Wu, Hongyu  
TITLE OF INVENTION: Immunostimulatory Compositions  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Denlinger & Associates  
STREET: 350 Cambridge Ave. Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/146, 283  
FILING DATE: 03-SEPT-1998  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Judge, Linda R.  
REGISTRATION NUMBER: 42,702  
REFERENCE/DOCKET NUMBER: 7636-0010.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0960  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 782 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: homo sapiens  
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
US-09-146-283-4

Query Match 42.3%; Score 292; DB 2; Length 782;  
Best Local Similarity 100.0%; Pred. No. 3.5e-269;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 38 GAATOVCTGTDMKRLRASPETHLMDLRHLXYGCGVQVGNLELTPTNLSLFLDIO 97  
Db 19 GAATOVCTGTDMKRLRASPETHLMDLRHLXYGCGVQVGNLELTPTNLSLFLDIO 78  
Qy 98 EVGGVLIANHNOVQVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 157  
Db 79 EVGGVLIANHNOVQVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 138

Qy 158 ELQLSRLEILKGVLIQIORNPOLCYODITLWKDIFHKNNQALTLIDITNRSRACHPCSPM 217  
Db 139 ELQLSRLEILKGVLIQIORNPOLCYODITLWKDIFHKNNQALTLIDITNRSRACHPCSPM 198  
Qy 218 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 277  
Db 199 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 258  
Qy 278 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNLTSDVGS 329  
Db 259 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNLTSDVGS 310

RESULT 3  
US-08-579-823A-4  
Sequence 4, Application US/08579823A  
Patent No. 6080409

GENERAL INFORMATION:  
APPLICANT: Laus, Reiner  
APPLICANT: Ruegg, Curtis L.  
APPLICANT: Wu, Hongyu  
TITLE OF INVENTION: Immunostimulatory Composition and Method  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Denlinger & Associates  
STREET: 350 Cambridge Ave. Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/579, 823A  
FILING DATE: 03-DEC-1998  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Judge, Linda R.  
REGISTRATION NUMBER: 42,702  
REFERENCE/DOCKET NUMBER: 7636-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0960  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 782 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: homo sapiens  
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
US-08-579-823A-4

Query Match 42.3%; Score 292; DB 3; Length 782;  
Best Local Similarity 100.0%; Pred. No. 3.5e-269;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 38 GAATOVCTGTDMKRLRASPETHLMDLRHLXYGCGVQVGNLELTPTNLSLFLDIO 97  
Db 19 GAATOVCTGTDMKRLRASPETHLMDLRHLXYGCGVQVGNLELTPTNLSLFLDIO 78  
Qy 98 EVGGVLIANHNOVQVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 157  
Db 79 EVGGVLIANHNOVQVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 138  
Qy 158 ELQLSRLEILKGVLIQIORNPOLCYODITLWKDIFHKNNQALTLIDITNRSRACHPCSPM 217  
Db 139 ELQLSRLEILKGVLIQIORNPOLCYODITLWKDIFHKNNQALTLIDITNRSRACHPCSPM 198

QY 218 CAGSRGWGESSSEDGSLRTVTCAGGACARCKGGLPTDCCHGCAAGCTGPKHSDCLACLHF 277  
DB 199 CAGSRGWGESSSEDGSLRTVTCAGGACARCKGGLPTDCCHGCAAGCTGPKHSDCLACLHF 258  
QY 278 NHSGICELHCPALVYNTDTEFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329  
DB 259 NHSGICELHCPALVYNTDTEFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310

## RESULT 4

US-09-344-195-4  
; Sequence 4, Application US/09344195  
; Patent No. 6210662

## GENERAL INFORMATION:

APPLICANT: Laus, Reiner  
Ruegg, Curtis L.  
Wu, Hongyu

TITLE OF INVENTION: Immunostimulatory Compositions

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave. Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/344,195

FILING DATE: 24-Jun-1999

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/146,283

FILING DATE: 03-SEPT-1998

ATTORNEY/AGENT INFORMATION:

NAME: Judge, Linda R.

REGISTRATION NUMBER: 42,702

REFERENCE/DOCKET NUMBER: 7636-0010.21

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-0880

TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 782 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORGANISM: homo sapiens

INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-344-195-4

Query Match 42.3%; Score 292; DB 4; Length 782;

Best Local Similarity 100.0%; Pred. No. 3.5e-269;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAASOVCTGTGTMKRLRPA SPETHLDMLRHLYOGGCOVVGNELELYLPNTNASTSLFLDIO 97  
DB 19 GAASOVCTGTGTMKRLRPA SPETHLDMLRHLYOGGCOVVGNELELYLPNTNASTSLFLDIO 78  
QY 98 EVQGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLR 157  
DB 79 EVQGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLR 138  
QY 158 ELQLSLTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQALATLIDTNRSAACHPCSPM 217

DB 139 ELQLSLTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQALATLIDTNRSAACHPCSPM 198  
QY 218 CAGSRGWGESSSEDGSLRTVTCAGGACARCKGGLPTDCCHGCAAGCTGPKHSDCLACLHF 277  
DB 199 CAGSRGWGESSSEDGSLRTVTCAGGACARCKGGLPTDCCHGCAAGCTGPKHSDCLACLHF 258  
QY 278 NHSGICELHCPALVYNTDTEFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329  
DB 259 NHSGICELHCPALVYNTDTEFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310

## RESULT 5

US-08-467-083-68  
; Sequence 68, Application US/08467083  
; Patent No. 5726023

## GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.  
Disis, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

NUMBER OF SEQUENCES: 68

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,083

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/414,417

FILING DATE: 06-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

TELEX: 3723836 SEDANBERRY

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-467-083-68

Query Match 42.3%; Score 292; DB 1; Length 1255;

Best Local Similarity 100.0%; Pred. No. 5.5e-269;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAASOVCTGTGTMKRLRPA SPETHLDMLRHLYOGGCOVVGNELELYLPNTNASTSLFLDIO 97  
DB 19 GAASOVCTGTGTMKRLRPA SPETHLDMLRHLYOGGCOVVGNELELYLPNTNASTSLFLDIO 78  
QY 98 EVQGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLR 157  
DB 79 EVQGYVLAHNOVROVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLR 138  
QY 158 ELQLSLTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQALATLIDTNRSAACHPCSPM 217  
DB 139 ELQLSLTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQALATLIDTNRSAACHPCSPM 198  
QY 218 CAGSRGWGESSSEDGSLRTVTCAGGACARCKGGLPTDCCHGCAAGCTGPKHSDCLACLHF 277

|||||  
Db 199 CKGRCHGESSBDCSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 258  
Qy 278 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSDVGS 329  
|||||  
Db 259 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSDVGS 310  
RESULT 6  
US-08-414-17B-68  
: Sequence 68, Application US/08414417B  
: Patent No. 5801005  
: GENERAL INFORMATION:  
: APPLICANT: Cheever, Martin A.  
: APPLICANT: Disis, Mary L.  
: TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
: TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
: TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
: NUMBER OF SEQUENCES: 69  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Seed and Berry LLP  
: STREET: 6300 Columbia Center, 701 Fifth Avenue  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: US  
: ZIP: 98104-7092  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: OPERATING SYSTEM: IBM PC compatible  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/414,417B  
: FILING DATE: 31-MAR-1995  
: CLASSIFICATION: 424  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Sharkey, Richard G.  
: REGISTRATION NUMBER: 32,629  
: REFERENCE/DOCKET NUMBER: 920010.448C2  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 622-4900  
: TELEFAX: (206) 682-6031  
: INFORMATION FOR SEQ ID NO: 68:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1255 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: US-08-414-17B-68  
Query Match 42.3%; Score 292; DB 1: Length 1255;  
Best Local Similarity 100.0%; Pred. No. 5.5e-269;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|||||  
Db 199 CKGRCHGESSBDCSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 258  
Qy 278 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSDVGS 329  
|||||  
Db 259 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSDVGS 310  
RESULT 7  
US-08-484-438-8  
: Sequence 8, Application US/08484438  
: Patent No. 5811098  
: Patent No. 5811098 5780031  
: GENERAL INFORMATION:  
: APPLICANT: PLOWMAN, Gregory D.  
: APPLICANT: Culouscou, Jean-Michel  
: APPLICANT: Shoyab, Mohammed  
: APPLICANT: Slegall, Clay B.  
: APPLICANT: Hellstr m, Ingegerd  
: APPLICANT: Hellstr m, Karl E.  
: TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
: NUMBER OF SEQUENCES: 42  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Pennile & Edmonds  
: STREET: 1155 Avenue of the Americas  
: CITY: New York  
: STATE: New York  
: COUNTRY: U.S.A.  
: ZIP: 10036-2711  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: OPERATING SYSTEM: IBM PC compatible  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/484,438  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 530  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/323,442  
: FILING DATE: 14-OCT-1994  
: APPLICATION NUMBER: US 08/150,704  
: FILING DATE: 10-NOV-1993  
: CLASSIFICATION: 530  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/981,165  
: FILING DATE: 24-NOV-1992  
: CLASSIFICATION: 530  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Mastrock, S. Leslie  
: REGISTRATION NUMBER: 18,872  
: REFERENCE/DOCKET NUMBER: 5624-230  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (212) 790-9090  
: TELEFAX: (212) 869-8864/9741  
: TELEX: 66141 PENNIE  
: INFORMATION FOR SEQ ID NO: 8:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1255 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: unknown  
: TOPOLOGY: unknown  
: MOLECULE TYPE: protein  
: US-08-484-438-8  
Query Match 42.3%; Score 292; DB 2: Length 1255;  
Best Local Similarity 100.0%; Pred. No. 5.5e-269;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 CKGSRCKWESSSEDCOSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 277  
| | | | |  
DB 199 CKGSRCKWESSSEDCOSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 258  
| | | | |  
OY 278 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329  
| | | | |  
DB 259 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310  
| | | | |

## RESULT 8

US-08-486-348A-68  
: Sequence 68, Application US/08486348A  
: Patent No. 5846538  
: GENERAL INFORMATION:  
: APPLICANT: Cheever, Martin A.  
: APPLICANT: Disis, Mary L.  
: TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
: TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
: TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
: NUMBER OF SEQUENCES: 69  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Seed and Berry LLP  
: STREET: 6300 Columbia Center, 701 Fifth Avenue  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: US  
: ZIP: 98104-7092  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/486,348A  
: FILING DATE: 07-JUN-1995  
: CLASSIFICATION: 424  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Sharkey, Richard G.  
: REGISTRATION NUMBER: 32,629  
: REFERENCE/DOCKET NUMBER: 920010.448C6  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 622-4900  
: TELEFAX: (206) 682-6031  
: INFORMATION FOR SEQ ID NO: 68:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1255 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
US-08-486-348A-68

Query Match 42.3%; Score 292; DB 2; Length 1255;

Best Local Similarity 100.0%; Pred. No. 5.5e-269; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAASOVCTGTDMLRLPASPETHDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 97  
| | | | |  
DB 19 GAASOVCTGTDMLRLPASPETHDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 78  
| | | | |  
OY 98 EVOGVYLLAHNOROVPLQRLRIVRGTOLEFEDNYALAVLNDNDPLNNTTPVYGASPGGLR 157  
| | | | |  
DB 79 EVOGVYLLAHNOROVPLQRLRIVRGTOLEFEDNYALAVLNDNDPLNNTTPVYGASPGGLR 138  
| | | | |  
OY 158 ELQLSLTEILKGVLIQNPOLCYODITLMDIFHKNNQALTLTIDNRSRACHPCSPM 217  
| | | | |  
DB 139 ELQLSLTEILKGVLIQNPOLCYODITLMDIFHKNNQALTLTIDNRSRACHPCSPM 198  
| | | | |  
OY 218 CKGSRCKWESSSEDCOSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 277  
| | | | |  
DB 199 CKGSRCKWESSSEDCOSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 258  
| | | | |  
OY 278 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329  
| | | | |

DB 259 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310  
| | | | |

## RESULT 9

US-08-625-101-2  
: Sequence 2, Application US/08625101  
: Patent No. 5869445  
: GENERAL INFORMATION:  
: APPLICANT: Cheever, Martin A.  
: APPLICANT: Disis, Mary L.  
: TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
: TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
: TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
: TITLE OF INVENTION: ONCOGENE IS ASSOCIATED  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: SEED AND BERRY LLP  
: STREET: 6300 Columbia Center, 701 Fifth Avenue  
: CITY: Seattle  
: STATE: Washington  
: COUNTRY: USA  
: ZIP: 98104-7092  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/625,101  
: FILING DATE: 01-APR-1996  
: CLASSIFICATION: 424  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Sharkey, Richard G.  
: REGISTRATION NUMBER: 32,629  
: REFERENCE/DOCKET NUMBER: 920010.448C7  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (206) 622-4900  
: TELEFAX: (206) 682-6031  
: INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 1255 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
US-08-625-101-2

Query Match 42.3%; Score 292; DB 2; Length 1255;

Best Local Similarity 100.0%; Pred. No. 5.5e-269; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAASOVCTGTDMLRLPASPETHDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 97  
| | | | |  
DB 19 GAASOVCTGTDMLRLPASPETHDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 78  
| | | | |  
OY 98 EVOGVYLLAHNOROVPLQRLRIVRGTOLEFEDNYALAVLNDNDPLNNTTPVYGASPGGLR 157  
| | | | |  
DB 79 EVOGVYLLAHNOROVPLQRLRIVRGTOLEFEDNYALAVLNDNDPLNNTTPVYGASPGGLR 138  
| | | | |  
OY 158 ELQLSLTEILKGVLIQNPOLCYODITLMDIFHKNNQALTLTIDNRSRACHPCSPM 217  
| | | | |  
DB 139 ELQLSLTEILKGVLIQNPOLCYODITLMDIFHKNNQALTLTIDNRSRACHPCSPM 198  
| | | | |  
OY 218 CKGSRCKWESSSEDCOSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 277  
| | | | |  
DB 199 CKGSRCKWESSSEDCOSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 258  
| | | | |  
OY 278 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329  
| | | | |  
DB 259 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310  
| | | | |  
RESULT 10  
US-08-468-545B-68

Sequence 68, Application US/08468545B  
Patent No. 5876712  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,545B  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 20010.448C5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-468-545B-68

Query Match 42.3%; Score 292; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 5.5e-269;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGDMKRLRPASPEHLDMLRHLHYCCQVVGMLLELYLPTNLSFLDDIQ 97  
DB 19 GAATGCTGDMKRLRPASPEHLDMLRHLHYCCQVVGMLLELYLPTNLSFLDDIQ 78  
QY 98 EVGGYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDDPLNNTPTVTGASPGGLR 157  
DB 79 EVGGYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDDPLNNTPTVTGASPGGLR 138  
QY 158 ELQRLSLTEILKGGVLIQRPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217  
DB 139 ELQRLSLTEILKGGVLIQRPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198  
QY 218 CKGRRCGSESEDCOSLTRFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACLHF 277  
DB 199 CKGRRCGSESEDCOSLTRFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACLHF 258  
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329  
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 11  
US-08-356-786-2  
Sequence 2, Application US/08356786  
Patent No. 5877305  
GENERAL INFORMATION:  
APPLICANT: Huston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.

TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
TITLE OF INVENTION: Marker  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,786  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-356-786-2

Query Match 42.3%; Score 292; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 5.5e-269;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGDMKRLRPASPEHLDMLRHLHYCCQVVGMLLELYLPTNLSFLDDIQ 97  
DB 19 GAATGCTGDMKRLRPASPEHLDMLRHLHYCCQVVGMLLELYLPTNLSFLDDIQ 78  
QY 98 EVGGYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDDPLNNTPTVTGASPGGLR 157  
DB 79 EVGGYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDDPLNNTPTVTGASPGGLR 138  
QY 158 ELQRLSLTEILKGGVLIQRPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217  
DB 139 ELQRLSLTEILKGGVLIQRPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198  
QY 218 CKGRRCGSESEDCOSLTRFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACLHF 277  
DB 199 CKGRRCGSESEDCOSLTRFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACLHF 258  
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329  
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 12  
US-08-466-680B-68  
Sequence 68, Application US/08466680B  
Patent No. 6075122  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 69

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

RESULT 14
US-08-422-734-1
Sequence 1, Application US/08422734
Patent No. 6333169
GENERAL INFORMATION:
APPLICANT: Huddiak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

```

```

; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genetech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,734
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/422108
; FILING DATE: 14-Apr-1995
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-422-734-1

Query Match 41.9%; Score 289; DB 4; Length 624;
Best Local Similarity 100.0%; Pred. No. 2e-266;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 STQCTGTDMKRLRPASPEHLDMLRHLYOGCQVVGKLELYLPTNASLSFLDDIOEVQ 100
DB 1 STQCTGTDMKRLRPASPEHLDMLRHLYOGCQVVGKLELYLPTNASLSFLDDIOEVQ 60
QY 101 GYVLIANQVQVPLQRLRIYRGTOLEFDNATALAVLDNGDPLNTPVTGASPGGLRELQ 160
DB 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNATALAVLDNGDPLNTPVTGASPGGLRELQ 120
QY 161 LRSLTILKGGVLIQNRNOLCYODTILMKDIFHNNOALATLIDTNSRACHPCSPMCKG 220
DB 121 LRSLTILKGGVLIQNRNOLCYODTILMKDIFHNNOALATLIDTNSRACHPCSPMCKG 180
QY 221 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 280
DB 181 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 240
QY 281 GICELHCPALVTYNTDIFESMPNPEGRTTGASCVTACPYNYLSTDVGS 329
DB 241 GICELHCPALVTYNTDIFESMPNPEGRTTGASCVTACPYNYLSTDVGS 289

RESULT 15
US-08-414-417B-69
; Sequence 69, Application US/08414417B
; Patent No. 5801005
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-414-417B-69

Query Match 19.1%; Score 132; DB 1; Length 580;
Best Local Similarity 100.0%; Pred. No. 4.1e-117;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 GAGGVMVHRRSSSTRSGGDLTLGLEPSEEPSPPLAPSEGASDVFDGLGMAAKG 389
DB 363 GAGGVMVHRRSSSTRSGGDLTLGLEPSEEPSPPLAPSEGASDVFDGLGMAAKG 422
QY 390 LQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVQPDVRRQPPSPREGPL 449
DB 423 LQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVQPDVRRQPPSPREGPL 482
QY 450 PAARPAGATTLER 461
DB 483 PAARPAGATTLER 494

Search completed: Apr-11 28, 2003, 13:46:38
Job time : 19 secs

```





```

Db 181 CYPOTILMKDIFHNKNOTALTLIDNRSRACHPCSPMKGRCWGESSEDCOSLTRVCA 240
Qy 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALTYNTDPRES 300
Db 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALTYNTDPRES 300
Qy 301 MPNEBGRITFGASCVTACPYNYLSTDVSGAGGWNHNRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNEBGRITFGASCVTACPYNYLSTDVSGAGGWNHNRSSSTRSGGDLTLGLEPSEE 360
Qy 361 EAPNSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQRYSEDPVLPSTDSYV 420
Db 361 EAPNSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQRYSEDPVLPSTDSYV 420
Qy 421 APLTCSPOPEYVNOPDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVVDVFAFG 480
Db 421 APLTCSPOPEYVNOPDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVVDVFAFG 480
Qy 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPAENPEY 540
Qy 541 LGLDVPAAPARSPSPQPPWEHNAIOEARLNLSPDTAAENNETYEVISEMFDLOEP 600
Db 541 LGLDVPAAPARSPSPQPPWEHNAIOEARLNLSPDTAAENNETYEVISEMFDLOEP 600
Qy 601 TCLQTRLELYKQGLRSLTKLKGPLTWMASHYKOHCPPTPESCATQIITFESEKENLD 660
Db 601 TCLQTRLELYKQGLRSLTKLKGPLTWMASHYKOHCPPTPESCATQIITFESEKENLD 660
Qy 661 FLVVIPEDCWEPVOGAPPPAAAHNNH 690
Db 661 FLVVIPEDCWEPVOGAPPPAAAHNNH 690

```

## RESULT 2

US-09-821-883-1

```

: Sequence 1, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Retner
: APPLICANT: Vidovic, Damir
: APPLICANT: Gradis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE OF INVENTION: Cell-Based Immunotherapy
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: CURRENT FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO. 1
: LENGTH: 555
: TYPE: PR
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500 construct
US-09-821-883-1

```

```

Query Match 79.6%; Score 549; DB 10; Length 555;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MRAAPLLARASLSLGLFLFFWLDRLSVLAKELARGAASSTOYCTGDMKRLRPASPE 60
Db 1 MRAAPLLARASLSLGLFLFFWLDRLSVLAKELARGAASSTOYCTGDMKRLRPASPE 60
Qy 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASLSFLDIDIOEVGYVLIANOVQVPLQRL 120
Db 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASLSFLDIDIOEVGYVLIANOVQVPLQRL 120

```

```

Qy 121 VRGQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLRELQRLSLTEILKGVLIOBNOL 180
Db 121 VRGQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLRELQRLSLTEILKGVLIOBNOL 180
Qy 181 CYPOTILMKDIFHNKNOTALTLIDNRSRACHPCSPMKGRCWGESSEDCOSLTRVCA 240
Db 181 CYPOTILMKDIFHNKNOTALTLIDNRSRACHPCSPMKGRCWGESSEDCOSLTRVCA 240
Qy 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALTYNTDPRES 300
Db 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALTYNTDPRES 300
Qy 301 MPNEBGRITFGASCVTACPYNYLSTDVSGAGGWNHNRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNEBGRITFGASCVTACPYNYLSTDVSGAGGWNHNRSSSTRSGGDLTLGLEPSEE 360
Qy 361 EAPNSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQRYSEDPVLPSTDSYV 420
Db 361 EAPNSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQRYSEDPVLPSTDSYV 420
Qy 421 APLTCSPOPEYVNOPDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVVDVFAFG 480
Db 421 APLTCSPOPEYVNOPDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVVDVFAFG 480
Qy 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPAENPEY 540
Qy 541 LGLDVPAAP 549
Db 541 LGLDVPAAP 549

```

## RESULT 3

US-09-821-883-5

```

: Sequence 5, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Retner
: APPLICANT: Vidovic, Damir
: APPLICANT: Gradis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE OF INVENTION: Cell-Based Immunotherapy
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: CURRENT FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO. 5
: LENGTH: 479
: TYPE: PR
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER300*-TGM-CSF construct
US-09-821-883-5

```

```

Query Match 47.7%; Score 329; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 4,1e-267;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MRAAPLLARASLSLGLFLFFWLDRLSVLAKELARGAASSTOYCTGDMKRLRPASPE 60
Db 1 MRAAPLLARASLSLGLFLFFWLDRLSVLAKELARGAASSTOYCTGDMKRLRPASPE 60
Qy 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASLSFLDIDIOEVGYVLIANOVQVPLQRL 120
Db 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASLSFLDIDIOEVGYVLIANOVQVPLQRL 120
Qy 121 VRGQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLRELQRLSLTEILKGVLIOBNOL 180
Db 121 VRGQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLRELQRLSLTEILKGVLIOBNOL 180

```

QY 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMKGSRMGESSEDCOSLTRTYCA 240  
Db 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMKGSRMGESSEDCOSLTRTYCA 240  
QY 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFES 300  
Db 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFES 300  
QY 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329  
Db 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329

## RESULT 4

US-09-821-883-3  
Sequence 3, Application US/09821883  
Patent No. US20020061310A1  
GENERAL INFORMATION:  
APPLICANT: Laus, Relner  
APPLICANT: Vidovic, Damir  
APPLICANT: Graddis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821,883  
CURRENT FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/193,504  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 564  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500\* construct  
US-09-821-883-3

Query Match 47.7%; Score 329; DB 10; Length 564;

Best Local Similarity 100.0%; Pred. No. 4.8e-267; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELARGAASVOCTGDMKRLPASPT 60  
Db 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELARGAASVOCTGDMKRLPASPT 60  
QY 61 HLDMLRHLVYOGCOVVOGNLELTYLPTNNSLSFLQDIOEVQGVYLLAHNOVROPVLRRLI 120  
Db 61 HLDMLRHLVYOGCOVVOGNLELTYLPTNNSLSFLQDIOEVQGVYLLAHNOVROPVLRRLI 120  
QY 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGGLRELQLSLEILLKGVLIORNPOL 180  
Db 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGGLRELQLSLEILLKGVLIORNPOL 180  
QY 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMKGSRMGESSEDCOSLTRTYCA 240  
Db 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMKGSRMGESSEDCOSLTRTYCA 240  
QY 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFES 300  
Db 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFES 300  
QY 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329  
Db 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329

## RESULT 5

US-09-821-883-4  
Sequence 4, Application US/09821883  
Patent No. US20020061310A1  
GENERAL INFORMATION:

APPLICANT: Laus, Relner  
APPLICANT: Vidovic, Damir  
APPLICANT: Graddis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821,883  
CURRENT FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/193,504  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 697  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500\*-IGM-CSF construct  
US-09-821-883-4

Query Match 47.7%; Score 329; DB 10; Length 697;

Best Local Similarity 100.0%; Pred. No. 5.7e-267; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELARGAASVOCTGDMKRLPASPT 60  
Db 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELARGAASVOCTGDMKRLPASPT 60  
QY 61 HLDMLRHLVYOGCOVVOGNLELTYLPTNNSLSFLQDIOEVQGVYLLAHNOVROPVLRRLI 120  
Db 61 HLDMLRHLVYOGCOVVOGNLELTYLPTNNSLSFLQDIOEVQGVYLLAHNOVROPVLRRLI 120  
QY 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGGLRELQLSLEILLKGVLIORNPOL 180  
Db 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGGLRELQLSLEILLKGVLIORNPOL 180  
QY 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMKGSRMGESSEDCOSLTRTYCA 240  
Db 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMKGSRMGESSEDCOSLTRTYCA 240  
QY 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFES 300  
Db 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTFES 300  
QY 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329  
Db 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329

RESULT 6  
US-09-921-161-1

Sequence 1, Application US/09921161  
Patent No. US20020090662A1  
GENERAL INFORMATION:  
APPLICANT: Ralph, Peter  
TITLE OF INVENTION: ANALYTICAL METHOD  
FILE REFERENCE: GENENT.066A  
CURRENT APPLICATION NUMBER: US/09/921,161  
CURRENT FILING DATE: 2001-08-01  
PRIOR APPLICATION NUMBER: 60/225,433  
PRIOR FILING DATE: 2000-08-15  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 645  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-921-161-1

Query Match 42.3%; Score 292; DB 10; Length 645;

Best Local Similarity 100.0%; Pred. No. 4.7e-226; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match	42.3%	Score 292	DB 9	Length 653
Best Local Similarity	100.0%	Pred. No. 4	Be-236	
Matches	292	Conservative	0	Mismatches 0; Indels 0; Gaps 0
QY	38	GAASQVCTGTDMKRLPASPETHLDMRLNLYGSCVVOGNIETLYLPTNALSLSFDIOQ	97	
Db	19	GAASQVCTGTDMKRLPASPETHLDMRLNLYGSCVVOGNIETLYLPTNALSLSFDIOQ	78	
QY	98	EVQGVYLLAHNNOVROPRLRLRYVGTOLEFDNVALAVLDNCDPLNNTPTVIGASPGGR	157	
Db	79	EVQGVYLLAHNNOVROPRLRLRYVGTOLEFDNVALAVLDNCDPLNNTPTVIGASPGGR	138	
QY	158	ELDLRLSTLEILKGYLLIORNPOLCYODTILMKDIFHKNNQALALFTLIDNRSFACHPCSPM	217	
Db	139	ELDLRLSTLEILKGYLLIORNPOLCYODTILMKDIFHKNNQALALFTLIDNRSFACHPCSPM	198	
QY	218	CKSGKMCESSESDCOSLRLTVACAGCARGKGLPTDCHEGCAACGTGPKHSDCLACHF	277	
Db	199	CKSGKMCESSESDCOSLRLTVACAGCARGKGLPTDCHEGCAACGTGPKHSDCLACHF	258	
QY	278	NHSGICELHACALVYNTDTEFSPMPNPGRTYFGASCATPCPYNTLSTDVGS	329	
Db	259	NHSGICELHACALVYNTDTEFSPMPNPGRTYFGASCATPCPYNTLSTDVGS	310	

Query Match	Similarity	42.3%	Score 292	DB 9	Length 712
Best Local	Similarity	100.0%	Pred. No. 5.1e-236		
Matches	Conservative	0	Mismatches	0	Indels
					Gaps
QY	38	GAASIVGTGTGTMKRLRLASPEETHDMLRNHYVGGCGVAGNINLETYLPPTNASTLSTLQDIO	97		
Db	19	GAASIVGTGTGTMKRLRLASPEETHDMLRNHYVGGCGVAGNINLETYLPPTNASTLSTLQDIO	78		
QY	98	EVQGVYLLAHNOVROVPLQRLRIYRGTOLEFDNVALAVLNDGDPINNTTPLYTGASPGGLR	157		
Db	79	EVQGVYLLAHNOVROVPLQRLRIYRGTOLEFDNVALAVLNDGDPINNTTPLYTGASPGGLR	138		
QY	158	ELQRLSLTEILKGGVLIQIRNPOLCYQDTITLKKDIFHKNNQALTLIDTNRSRACIRPCSM	217		
Db	139	ELQRLSLTEILKGGVLIQIRNPOLCYQDTITLKKDIFHKNNQALTLIDTNRSRACIRPCSM	198		
QY	218	CKGSRCKMESSSDQSLRTRYVAGGCACRKGRLPRDCCHEQCAACSTGRKSIDCLACHF	277		
Db	199	CKGSRCKMESSSDQSLRTRYVAGGCACRKGRLPRDCCHEQCAACSTGRKSIDCLACHF	258		
QY	278	NHSGICEIHCPLALVYNTDTFESMNPFGGRYTFGASCTYACPYNTLSTDVGS	329		
Db	239	NHSGICEIHCPLALVYNTDTFESMNPFGGRYTFGASCTYACPYNTLSTDVGS	310		

RESULT 5  
 US-09-854-356-6  
 Sequence 6, Application US/09854356  
 Patent No. US20020177567A1  
 GENERAL INFORMATION:  
 APPLICANT: Cheever, Martin A.  
 APPLICANT: Chaysen, Dirk  
 APPLICANT: Corixa Corporation  
 APPLICANT: SmithKline Beecham Biologicals S. A.  
 TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 FILE REFERENCE: 014058-009810PC  
 CURRENT APPLICATION NUMBER: US/09/854.356  
 CURRENT FILING DATE: 2001-05-09  
 PRIOR APPLICATION NUMBER: US 09/493,480  
 PRIOR FILING DATE: 2000-01-28  
 PRIOR APPLICATION NUMBER: US 60/117,976  
 PRIOR FILING DATE: 1999-01-29  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 6



RESULT 12  
 US-09-930-125-2  
 : Sequence 2, Application US/09930125  
 : Publication No. US20020193329A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Hand-Zimmerman, Susan  
 : APPLICANT: Cheever, Martin A.  
 : APPLICANT: Foy, Teresa M.  
 : APPLICANT: Lodes, Michael J.  
 : APPLICANT: Kalos, Michael D.  
 : APPLICANT: McNeill, Patricia D.  
 : APPLICANT: Vegdick, Thomas S.  
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
 : TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
 : FILE REFERENCE: 210121.544  
 : CURRENT APPLICATION NUMBER: US/09/930.125  
 : CURRENT FILING DATE: 2001-08-14  
 : NUMBER OF SEQ ID NOS: 25  
 : SOFTWARE: FastSeq for Windows Version 3.0  
 : SEQ ID NO 2  
 : LENGTH: 1255  
 : TYPE: PRT  
 : ORGANISM: Homo sapien  
 US-09-930-125-2

Query Match 42.3%; Score 292; DB 9; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-236;  
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTLDIO 97  
 DB 19 GAATGCTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTLDIO 78  
 QY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 157  
 DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 138  
 QY 158 ELQRLSLLEILKGVLIORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217  
 DB 139 ELQRLSLLEILKGVLIORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198  
 QY 218 CKGRMCWESSEDCOSLRTVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLACHF 277  
 DB 199 CKGRMCWESSEDCOSLRTVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLACHF 258  
 QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329  
 DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 13  
 US-09-441-411-6  
 : Sequence 6, Application US/09441411  
 : Publication No. US20030008342A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Scholler, Nathalie B.  
 : APPLICANT: Disis, Mary L.  
 : APPLICANT: Hellicrom, Ingegerd  
 : APPLICANT: Hellicrom, Karl Erik  
 : TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
 : FILE REFERENCE: 730033.409  
 : CURRENT APPLICATION NUMBER: US/09/441.411  
 : CURRENT FILING DATE: 1999-11-16  
 : NUMBER OF SEQ ID NOS: 26  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 6  
 : LENGTH: 1255  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 US-09-441-411-6

Query Match 42.3%; Score 292; DB 9; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-236;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTLDIO 97  
 DB 19 GAATGCTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTLDIO 78  
 QY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 157  
 DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 138  
 QY 158 ELQRLSLLEILKGVLIORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217  
 DB 139 ELQRLSLLEILKGVLIORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198  
 QY 218 CKGRMCWESSEDCOSLRTVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLACHF 277  
 DB 199 CKGRMCWESSEDCOSLRTVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLACHF 258  
 QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329  
 DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 14  
 US-09-811-123-9  
 : Sequence 9, Application US/09811123  
 : Patent No. US20020001587A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Sharon Erickson  
 : APPLICANT: Ralph Schwall  
 : APPLICANT: Mark Sliwowski  
 : TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ETBB  
 : TITLE OF INVENTION: ANTIBODY-MAINTAINING NOID CONJUGATES  
 : FILE REFERENCE: GENE 073A2  
 : CURRENT APPLICATION NUMBER: US/09/811.123  
 : CURRENT FILING DATE: 2001-03-16  
 : PRIOR APPLICATION NUMBER: 60/238,327  
 : PRIOR FILING DATE: 2000-10-05  
 : PRIOR APPLICATION NUMBER: 09/602,530  
 : PRIOR FILING DATE: 2000-06-23  
 : NUMBER OF SEQ ID NOS: 11  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 9  
 : LENGTH: 1255  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 US-09-811-123-9

Query Match 42.3%; Score 292; DB 10; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-236;  
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTLDIO 97  
 DB 19 GAATGCTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTLDIO 78  
 QY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 157  
 DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDNGDPLNNTPTVTGASPGGLR 138  
 QY 158 ELQRLSLLEILKGVLIORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217  
 DB 139 ELQRLSLLEILKGVLIORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198  
 QY 218 CKGRMCWESSEDCOSLRTVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLACHF 277  
 DB 199 CKGRMCWESSEDCOSLRTVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLACHF 258  
 QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329  
 DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 15  
US-09-811-115-3  
: Sequence 3, Application US/09811115  
: Patent No. US20020035736A1  
: GENERAL INFORMATION:  
: APPLICANT: Erickson, Sharon  
: APPLICANT: Schwall, Ralph  
: APPLICANT: King, Kathleen  
: TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
: FILE REFERENCE: GENENT 034A  
: CURRENT APPLICATION NUMBER: US/09/811,115  
: CURRENT FILING DATE: 2001-03-16  
: PRIOR APPLICATION NUMBER: 60/189,844  
: PRIOR FILING DATE: 2000-03-16  
: NUMBER OF SEQ ID NOS: 4  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 3  
: LENGTH: 1255  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
US-09-811-115-3

Query Match 42.38; Score 292; DB 10; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 8.4e-236;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 38 GAATGCTGTDMKRLPASPETHDMLRHLYGCGVVGNNLELYLPTNASTSLFDIO 97  
DB 19 GAATGCTGTDMKRLPASPETHDMLRHLYGCGVVGNNLELYLPTNASTSLFDIO 78  
QY 98 EVGGYVLLAHNQVRQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLR 157  
DB 79 EVGGYVLLAHNQVRQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLR 138  
QY 158 ELQLRSLLEILKGGVLIQRNQLCYQDTILMKDIFHKNNQALATLIDITNRSRACHPCSPM 217  
DB 139 ELQLRSLLEILKGGVLIQRNQLCYQDTILMKDIFHKNNQALATLIDITNRSRACHPCSPM 198  
QY 218 CKGRMCWSESEDCSLRTVCAGGACARCKGPLPTDCCHECOCAGCTGPKHSDCLACLHF 277  
DB 199 CKGRMCWSESEDCSLRTVCAGGACARCKGPLPTDCCHECOCAGCTGPKHSDCLACLHF 258  
QY 278 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329  
DB 259 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

Search completed: April 28, 2003, 13:47:06  
Job time : 24 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:08 ; Search time 23.1973 Seconds  
(without alignments)  
1660.081 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STGVCTGTDKRLRLPASPT.....FGASCVTACPYNTLSTDVGS 289

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /SID52/gcgdata/geneseq-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	289	AAE13120	Mature human HER-2
2	1587	100.0	479	AAE13112	Human HER300-TCM-C
3	1587	100.0	555	AAE13108	Human HER500 fusio
4	1587	100.0	564	AAE13110	Human HER500 oncop
5	1587	100.0	645	AAE13108	Human ERBB2 extrac
6	1587	100.0	645	AAE13108	Human ERBB2 extrac
7	1587	100.0	653	AAE13100	Human HER-2/neu on
8	1587	100.0	653	AAE13100	Human HER500-TCM-C
9	1587	100.0	690	AAE13109	Human HER500-TCM-C
10	1587	100.0	697	AAE13111	Human HER500-TCM-C

11	1587	100.0	712	21	AAE13120	Human HER-2/neu fu
12	1587	100.0	712	23	AAE13149	Her-2/neu extracel
13	1587	100.0	782	18	AAE19764	Her-2/neu extracel
14	1587	100.0	919	21	AAE13120	Human HER-2/neu fu
15	1587	100.0	919	21	AAE13148	Human HER-2/neu fu
16	1587	100.0	951	21	AAE14993	DC8acfv-erbB2C fu
17	1587	100.0	1200	21	AAE13120	Human HER-2/neu pr
18	1587	100.0	1255	17	AAE13111	Human HER-2/neu pr
19	1587	100.0	1255	20	AAE13146	Human HER-2/neu on
20	1587	100.0	1255	21	AAE13198	Human HER-2/neu on
21	1587	100.0	1255	21	AAE14780	Human HER-2/neu pr
22	1587	100.0	1255	21	AAE14780	Human HER-2/neu pr
23	1587	100.0	1255	22	AAE12130	Human tyrosine kin
24	1587	100.0	1255	22	AAE12130	Human tyrosine kin
25	1587	100.0	1255	22	AAE12130	Human tyrosine kin
26	1587	100.0	1255	22	AAE12130	Human tyrosine kin
27	1587	100.0	1255	23	AAE12130	Human tyrosine kin
28	1587	100.0	1255	23	AAE12130	Human tyrosine kin
29	1587	100.0	1255	23	AAE12130	Human tyrosine kin
30	1587	100.0	1255	23	AAE12130	Human tyrosine kin
31	1587	100.0	1255	23	AAE12130	Human tyrosine kin
32	1571	99.0	1433	14	AAE13568	Sequence of c-erbB
33	1552	97.8	624	11	AAE13568	Extracellular port
34	1538	96.9	419	23	AAE13568	Human truncated HE
35	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
36	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
37	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
38	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
39	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
40	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
41	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
42	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
43	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
44	1536	96.8	419	22	AAE13568	Human p68HER-2 gen
45	1536	96.8	419	22	AAE13568	Human p68HER-2 gen

## ALIGNMENTS

RESULT 1	AAE13120	standard; Protein: 289 AA.
AC	AAE13120;	
XX	28-JAN-2002 (first entry)	
DE	Mature human HER-2 membrane distal extracellular domain.	
XX		
KW	Immunostimulatory fusion protein; IFP; antigen component; therapy;	
KW	immunostimulatory component; T-cell mediated immune response; DC;	
KW	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;	
KW	human; HER-2 membrane distal extracellular domain.	
OS	Homo sapiens.	
XX		
PN	WO200174855 A2.	
XX		
PD	11-OCT-2001.	
XX		
PF	30-MAR-2001; 2001WO-US10515.	
XX		
PR	30-MAR-2000; 2000US-193504P.	
XX		
PA	(DEND-) DENDREON CORP.	
XX		
PI	Laus R, Vidovic D, Graddis T;	
XX		
DR	WPI: 2001-662965/76.	
XX	N-PSDB: AAD21571.	
PT	An immunostimulatory fusion protein comprising the intracellular domain	

PT of HER-2 and an antigen elicits an immune response to the antigen and  
 is useful for the treatment of associated cancer associated -  
 Claim 6; Page 30; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is  
 CC mature human HER-2 membrane distal extracellular domain.  
 CC This sequence is used in the HER500 and HER300 GM-CSF fusion constructs  
 of the invention.

SO Sequence 289 AA:

Query Match 100.0%; Score 1587; DB 22; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-132;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKRLRPASPEHLDMLRLHLYGCGVVGNNLELTYPNTASLSFLDDIDDEVQ 60  
 DB 1 STQVCTGDMKRLRPASPEHLDMLRLHLYGCGVVGNNLELTYPNTASLSFLDDIDDEVQ 60  
 QY 61 GYVLIANNOVROVPLQRLIRVGTQLEFEDNVAVLAVDNGDPLNNTTPTVGTASPGGLRELO 120  
 DB 61 GYVLIANNOVROVPLQRLIRVGTQLEFEDNVAVLAVDNGDPLNNTTPTVGTASPGGLRELO 120  
 QY 121 LRSLTLEIKGVLQIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 180  
 DB 121 LRSLTLEIKGVLQIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 180  
 QY 181 SRCWGESSEDCQSLTRIVCAGGCARCKGRLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 240  
 DB 181 SRCWGESSEDCQSLTRIVCAGGCARCKGRLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 240  
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 289  
 DB 241 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 289

# RESULT 2

AAE13112 ID AAE13112 standard; Protein; 479 AA.

AC AAE13112;

DT 28-JAN-2002 (first entry)

DE Human HER300-rGM-CSF fusion construct comprising OVA-derived peptide.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW Immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;  
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;  
 KW ovalbumin-derived octapeptide; OVA; rat; HER300-rGM-CSF fusion protein.

OS Chimeric - Homo sapiens.

OS Chimeric - Rattus norvegicus.

OS Chimeric - Unidentified.

PN WO200174855-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001MO-US10515.  
 PR 30-MAR-2000; 2000US-193504P.

XX (DEND-) DENDREON CORP.  
 PA Laus: R, Vidovic D, Graddis T;  
 XX WPI; 2001-662965/76.  
 DR N-PSDB; AAD21568.

PT An immunostimulatory fusion protein comprising the intracellular domain  
 of HER-2 and an antigen elicits an immune response to the antigen and  
 is useful for the treatment of associated cancer associated -  
 Example 1; Page 27; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER300  
 CC rGM-CSF fusion protein construct which comprises human PAP  
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2  
 CC signal sequence, mature HER-2 membrane distal extracellular domain,  
 CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,  
 CC a rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence  
 and a C-terminal tag.

SO Sequence 479 AA:

Query Match 100.0%; Score 1587; DB 22; Length 479;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-131;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKRLRPASPEHLDMLRLHLYGCGVVGNNLELTYPNTASLSFLDDIDDEVQ 60  
 DB 41 STQVCTGDMKRLRPASPEHLDMLRLHLYGCGVVGNNLELTYPNTASLSFLDDIDDEVQ 100  
 QY 61 GYVLIANNOVROVPLQRLIRVGTQLEFEDNVAVLAVDNGDPLNNTTPTVGTASPGGLRELO 120  
 DB 101 GYVLIANNOVROVPLQRLIRVGTQLEFEDNVAVLAVDNGDPLNNTTPTVGTASPGGLRELO 160  
 QY 121 LRSLTLEIKGVLQIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 180  
 DB 161 LRSLTLEIKGVLQIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 220  
 QY 181 SRCWGESSEDCQSLTRIVCAGGCARCKGRLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 240  
 DB 221 SRCWGESSEDCQSLTRIVCAGGCARCKGRLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 280  
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 289  
 DB 281 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 329

# RESULT 3

AAE13108 ID AAE13108 standard; Protein; 555 AA.

AC AAE13108;

DT 28-JAN-2002 (first entry)

DE Human HER500 fusion protein construct.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW Immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human;  
 KW HER-2 protein; HER500 fusion protein.

XX OS Chimeric - Homo sapiens.  
XX OS Chimeric - Synthetic.  
XX PN WO200174855-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US10515.  
XX PR 30-MAR-2000; 2000US-193504P.  
XX PA (DEND-) DENDREON CORP.  
XX PI Laus R, Vidovic D, Graddis T;  
XX DR WPI; 2001-662965/76.  
XX DR N-PSDB; AAD21564.  
XX PT An immunostimulatory fusion protein comprising the intracellular domain  
XX PT of HER-2 and an antigen elicits an immune response to the antigen and  
XX PT is useful for the treatment of associated cancer associated -  
XX PS Claim 7; Page 26; 59pp; English.  
XX CC The invention relates to immunostimulatory fusion proteins (IFP) and  
XX CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
XX CC polypeptide antigen component and an immunostimulatory component derived  
XX CC from the intracellular domain of HER-2 protein which is effective to  
XX CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
XX CC immune response to the antigen. IFP or superactivated dendritic cells  
XX CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
XX CC associated with a particularly antigen. The present sequence is HER500  
XX CC fusion protein construct which comprises human PAP signal  
XX CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal  
XX CC sequence, mature HER-2 membrane distal extracellular and intracellular  
XX CC domains and a C-terminal tag.  
XX SQ Sequence 555 AA;  
Query Match 100.0%; Score 1587; DB 22; Length 555;  
Best Local Similarity 100.0%; Pred. No. 1.8e-131;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 60  
DB 41 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 100  
QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELQ 120  
DB 101 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELQ 160  
QY 121 LRSLEILKGGVLIQORNQOLCYQDTIIMKDIFFHKNQALATLIDTNSRACHPCSPMKG 180  
DB 161 LRSLEILKGGVLIQORNQOLCYQDTIIMKDIFFHKNQALATLIDTNSRACHPCSPMKG 220  
QY 181 SRCWGESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 240  
DB 221 SRCWGESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 280  
QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289  
DB 281 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329  
RESULT 4  
AAE13110  
ID AAE13110 standard; Protein: 564 AA.  
XX AC AAE13110;  
XX DT 28-JAN-2002 (first entry)  
XX

DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.  
XX  
XX  
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
KW immunostimulatory component; T-cell mediated immune response; DC;  
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
KW membrane distal intracellular domain; C-terminal tag; human; OVA;  
KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein..  
XX  
XX OS Chimeric - Homo sapiens.  
XX OS Chimeric - Unidentified.  
XX PN WO200174855-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US10515.  
XX PR 30-MAR-2000; 2000US-193504P.  
XX PA (DEND-) DENDREON CORP.  
XX PI Laus R, Vidovic D, Graddis T;  
XX DR WPI; 2001-662965/76.  
XX DR N-PSDB; AAD21566.  
XX PT An immunostimulatory fusion protein comprising the intracellular domain  
XX PT of HER-2 and an antigen elicits an immune response to the antigen and  
XX PT is useful for the treatment of associated cancer associated -  
XX PS Claim 7; Page 26; 59pp; English.  
XX CC The invention relates to immunostimulatory fusion proteins (IFP) and  
XX CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
XX CC polypeptide antigen component and an immunostimulatory component derived  
XX CC from the intracellular domain of HER-2 protein which is effective to  
XX CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
XX CC immune response to the antigen. IFP or superactivated dendritic cells  
XX CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
XX CC associated with a particularly antigen. The present sequence is HER500  
XX CC fusion protein construct which comprises human PAP signal  
XX CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal  
XX CC sequence, mature HER-2 membrane distal extracellular domain, an  
XX CC Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,  
XX CC HER-2 membrane distal intracellular domain and a C-terminal tag.  
XX SQ Sequence 564 AA;  
Query Match 100.0%; Score 1587; DB 22; Length 564;  
Best Local Similarity 100.0%; Pred. No. 1.9e-131;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 80  
DB 41 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 100  
QY 121 LRSLEILKGGVLIQORNQOLCYQDTIIMKDIFFHKNQALATLIDTNSRACHPCSPMKG 180  
DB 161 LRSLEILKGGVLIQORNQOLCYQDTIIMKDIFFHKNQALATLIDTNSRACHPCSPMKG 220  
QY 181 SRCWGESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 240  
DB 221 SRCWGESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 280  
QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289  
DB 281 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329

```

RESULT 5
AAB60408
ID AAB60408 standard; protein; 645 AA.
XX
AC AAB60408;
XX
DT 24-APR-2001 (first entry)
XX
DE Human ErbB2 oncoprotein, SEQ ID NO:13.
XX
KW Anti-ErbB2 monoclonal antibody 2C4; HER2; mouse; murine; humanised; VL:
KW light chain variable region; cancer; cytostatic; EGFR-expressing cancer;
KW epidermal growth factor receptor; colon cancer; rectal cancer; tumour;
KW colorectal cancer; non-small cell lung cancer; metastatic breast cancer;
KW affinity purification.
XX
OS Homo sapiens.
XX
PN WO200100245-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-US17366.
XX
PR 25-JUN-1999; 99US-0141316.
XX
PA (GETH ) GENENTECH INC.
XX
PI Adams CW, Presta LG, Sliwkowsky M;
XX
DR WPI: 2001-080862/09.
XX
PT Treating cancer in a human, where the cancer expresses epidermal growth
PT factor receptor (EGFR), comprises administering an antibody which binds
PT ErbB2.
XX
PS Example 1; Fig 1A; 89pp; English.
XX
CC The invention relates to a method for treating cancer in a human patient,
CC wherein the cancer expresses epidermal growth factor receptor (EGFR),
CC comprising administering an antibody which binds ErbB2 (HER2; AAB60408).
CC In particular, the anti-ErbB2 antibody is the murine monoclonal antibody
CC 2C4 (AAB60396, AAB60397) or a humanised version of 2C4 (AAB60398),
CC encoding a humanised ErbB2-binding antibody; vectors and host cells
CC comprising such nucleic acids; the recombinant production of a humanised
CC ErbB2-binding antibody; and an immunconjugate comprising a humanised
CC ErbB2-binding antibody and a cytotoxic drug. The ErbB2-binding antibodies
CC act by antagonising ErbB receptors, and as inhibitors of transforming
CC growth factor alpha (TGF-alpha)-activated mitogen activated protein
CC kinase (MAPK). The method of the invention is used for treating cancer,
CC especially colon cancer, rectal cancer, colorectal cancer, lung cancer,
CC (especially non-small cell lung cancer), or breast cancer (especially
CC metastatic breast cancer). The antibodies may also have non-therapeutic
CC uses e.g., as affinity purification agents. Using an antibody which binds
CC to ErbB2 to treat cancer is preferable to the use of EGFR-targeted
CC drugs, as EGFR is also highly expressed in other tissues such as the
CC liver and skin, where the active drug will also bind, with skin toxicity.
CC having been observed for EGFR-targeted drugs. Antibodies which bind
CC ErbB2 are anticipated to have a better safety profile than such drugs.
CC The present sequence represents human ErbB2.
XX
SQ Sequence 645 AA;

```

```

Query Match 100.0%; Score 1587; DB 22; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STGVCTGTDMKLRPAPEPHLDMRLHLYGCGVVGNGNLELTLYLPNASTLSPFDIOEYQ 60
DB 22 STGVCTGTDMKLRPAPEPHLDMRLHLYGCGVVGNGNLELTLYLPNASTLSPFDIOEYQ 81

```

```

QY 61 GYVLIHNOVROYPLORLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPYTGASRGLEIQ 120
DB 82 GYVLIHNOVROYPLORLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPYTGASRGLEIQ 141
QY 121 LRSLETLKGVLIORNPOLCYODTILMKDIFPKNNOLATLIDTNRSRCHCSPCKG 180
DB 142 LRSLETLKGVLIORNPOLCYODTILMKDIFPKNNOLATLIDTNRSRCHCSPCKG 201
QY 181 SRCWGESSEDCQSLTRTYVACGACRCARGLPTDCHEQCAAGCTGPKHSDCLAFHNHS 240
DB 202 SRCWGESSEDCQSLTRTYVACGACRCARGLPTDCHEQCAAGCTGPKHSDCLAFHNHS 261
QY 241 GIEELHCPALVYNTDTFESMPPEGRYRTGASCYVACPNYSTDVG 289
DB 262 GIEELHCPALVYNTDTFESMPPEGRYRTGASCYVACPNYSTDVG 310

```

```

RESULT 6
AAB61593
ID AAB61593 standard; protein; 645 AA.
XX
AC AAB61593;
XX
DT 04-APR-2001 (first entry)
XX
DE Human ErbB2 extracellular domain.
XX
KW Human; ErbB2; cytostatic; prostate cancer; receptor tyrosine kinase;
KW antibody; ErbB receptor; monoclonal antibody 2C4; variable light chain.
XX
OS Homo sapiens.
XX
PN WO200100238-A1.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-US17423.
XX
PR 25-JUN-1999; 99US-0141315.
XX
PA (GETH ) GENENTECH INC.
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
PI Agus DB, Scher HT, Sliwkowski MX;
XX
DR WPI: 2001-159131/16.
XX
PT Treating prostate cancer in a human comprises administering an antibody
PT which binds ErbB2 and blocks ligand activation of an ErbB receptor.
XX
PS Disclosure; Fig 1; 93pp; English.
XX
CC The ErbB family of receptor tyrosine kinases are important mediators of
CC cell growth, differentiation and survival. The receptor family includes
CC four distinct members including Epidermal Growth Factor Receptor (EGFR or
CC ErbB1), HER2 (ErbB2 or p185neu), HER3 (ErbB3) and Her4 (ErbB4 or tyro2).
CC The present invention relates to a method for treating prostate cancer.
CC The method comprises administering an antibody which binds ErbB2 and
CC blocks ligand activation of an ErbB receptor. Preferably, the antibody
CC blocks binding of monoclonal antibody 2C4 to ErbB2 and/or blocks
CC TGF-alpha activation of mitogen-activated protein kinase (MAPK). The
CC present sequence is the extracellular domain of human ErbB2.
XX
SQ Sequence 645 AA;

```

```

Query Match 100.0%; Score 1587; DB 22; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STGVCTGTDMKLRPAPEPHLDMRLHLYGCGVVGNGNLELTLYLPNASTLSPFDIOEYQ 60
DB 22 STGVCTGTDMKLRPAPEPHLDMRLHLYGCGVVGNGNLELTLYLPNASTLSPFDIOEYQ 81

```

QY 61 GYVLIANOVQVPLQRLRIYRGTOLEFEDNYALVNDGDDPLNNTPTVTGASPGGLRELQ 120  
 DB 82 GYVLIANOVQVPLQRLRIYRGTOLEFEDNYALVNDGDDPLNNTPTVTGASPGGLRELQ 141  
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNNQALTLIDITNRSRACHPCSPMKCG 180  
 DB 142 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNNQALTLIDITNRSRACHPCSPMKCG 201  
 QY 181 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240  
 DB 202 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261  
 QY 241 GICELHCPALVTYNTDFTFESMPNEGRTFGASCVTACPYNYLSTDVGS 289  
 DB 262 GICELHCPALVTYNTDFTFESMPNEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 7  
 AAB21200  
 ID AAB21200 standard; protein: 653 AA.  
 AC AAB21200;  
 DT 12-JAN-2001 (first entry)  
 DE Extracellular HER-2/neu protein.  
 KW HER-2/neu: oncogene; tyrosine kinase; cytostatic; vaccine;  
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KW colon cancer.  
 OS Unidentified.  
 PN WO200044899-A1.  
 PD 03-AUG-2000.  
 PF 28-JAN-2000; 2000WO-US02164.  
 PR 29-JAN-1999; 99US-0117976.  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM.  
 PI Cheever MA, Gheysen D;  
 DR WPI: 2000-505976/45.  
 DR HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
 PT useful for vaccinating against breast, ovarian, colon, lung and  
 PT prostate cancers -  
 PS Claim 2: Fig 9; 128pp; English.  
 XX The present sequence is the extracellular HER-2/neu protein. HER-2/neu is  
 CC a member of the tyrosine kinase family of receptor-like glycoproteins and  
 CC shows homology to the epidermal growth factor receptor (EGFR). It  
 CC probably plays a part in cell growth and/or differentiation. The  
 CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising a  
 CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation  
 CC domain may be used to treat or prevent cancer by eliciting or  
 CC enhancing an immune response to the HER-2/neu protein. It may be used  
 CC to treat malignancies such as breast, ovarian, colon, lung and  
 CC prostate cancers, and may be used as an antigen to vaccinate against  
 CC these neoplasias.  
 XX  
 SQ Sequence 653 AA;  
 Query Match 100.0%; Score 1587; DB 21; Length 653;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-131;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 STOVCTGDMKRLRPASPEFHLDMLRHLXOCQVVGNLLETLYPTNASLSFLQDIOEVQ 60

DB 22 STOVCTGDMKRLRPASPEFHLDMLRHLXOCQVVGNLLETLYPTNASLSFLQDIOEVQ 81  
 QY 61 GYVLIANOVQVPLQRLRIYRGTOLEFEDNYALVNDGDDPLNNTPTVTGASPGGLRELQ 120  
 DB 82 GYVLIANOVQVPLQRLRIYRGTOLEFEDNYALVNDGDDPLNNTPTVTGASPGGLRELQ 141  
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNNQALTLIDITNRSRACHPCSPMKCG 180  
 DB 142 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNNQALTLIDITNRSRACHPCSPMKCG 201  
 QY 181 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240  
 DB 202 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261  
 QY 241 GICELHCPALVTYNTDFTFESMPNEGRTFGASCVTACPYNYLSTDVGS 289  
 DB 262 GICELHCPALVTYNTDFTFESMPNEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 8  
 AAM51145  
 ID AAM51145 standard; protein: 653 AA.  
 AC AAM51145;  
 DT 17-JUN-2002 (first entry)  
 DE Human Her-2/neu oncoprotein extracellular domain.  
 KW Her-2/neu: oncogene; cancer; tumour; vaccine; human; p185;  
 KW tyrosine kinase; receptor; c-erbB2; gene therapy.  
 OS Homo sapiens.  
 PN WO200212341-A2.  
 PD 14-FEB-2002.  
 PF 03-AUG-2001; 2001WO-US24283.  
 PR 03-AUG-2000; 2000US-0632507.  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.  
 PI Cheever MA, Gheysen D;  
 DR WPI: 2002-241743/29.  
 DR Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT or enhancing an immune response to the protein, has Her-2/neu  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain -  
 PS Claim 2: Fig 9; 141pp; English.  
 XX The present sequence is that of the extracellular domain of  
 CC human Her-2/neu (p185 glycoprotein or c-erbB2), an oncogenic  
 CC self-protein and target for anti-cancer vaccines. The Her-2/neu  
 CC gene is amplified and p185 is overexpressed in a variety of cancers,  
 CC including breast, ovarian, colon, lung and prostate cancer.  
 CC Her-2/neu (see AAM51143) is a member of the tyrosine kinase family  
 CC of receptor-like glycoproteins. It comprises an extracellular  
 CC domain with homology to the epidermal growth factor receptor  
 CC (EGFR), a highly hydrophobic transmembrane domain and a C-terminal  
 CC intracellular domain that also shows homology to EGFR. Its  
 CC overexpression correlates with a poor prognosis in breast and  
 CC ovarian cancers. The invention provides Her-2/neu fusion  
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines  
 CC comprising the fusion proteins or nucleic acid molecules. In  
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu  
 CC protein is fused to a Her-2/neu intracellular domain or

CC phosphorylation domain (or its DeltapD fragment). An immune  
 CC response to Her-2/neu protein is elicited or enhanced by  
 CC administering the fusion protein in the form of a vaccine, or by  
 CC transfecting cells of an animal *ex vivo* with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu  
 CC fusion protein can be used to remove tumour cells from a sample in  
 CC order to inhibit the development of cancer in a patient.

XX Sequence 653 AA:

Query Match 100.0%; Score 1587; DB 23; Length 653;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-131;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELTPLPNASLSFLQDIOEVQ 60  
 DB 22 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELTPLPNASLSFLQDIOEVQ 81  
 QY 61 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVPGASPGGLREIQ 120  
 DB 82 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVPGASPGGLREIQ 141  
 QY 121 LRSLTETLKGVLIQRNPOLCYODTILMKDIFRKNNOALTLTIDTNSRACHCSPCKG 180  
 DB 142 LRSLTETLKGVLIQRNPOLCYODTILMKDIFRKNNOALTLTIDTNSRACHCSPCKG 201  
 QY 181 SRGWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 240  
 DB 202 SRGWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 261  
 QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCATACPNYVLTSDVGS 289  
 DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCATACPNYVLTSDVGS 310

RESULT 9  
 AAE13109

ID AAE13109 standard; protein; 690 AA.

XX AAE13109;

DT 28-JAN-2002 (first entry)

DE Human HER500-hGM-CSF fusion protein construct.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW Immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;  
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;  
 KW HER500-hGM-CSF fusion protein.

XX Chimeric - Homo sapiens.

OS Chimeric - Synthetic.

XX MO200174855-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10515.

XX 30-MAR-2000; 2000US-193504P.

XX (DEND-) DENDREON CORP.

XX Laus R, Vldovic D, Graddis T;

XX WPI; 2001-662965/76.

DR N-PSDB; AAD21565.

XX An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX Claim 7; Page 26; 59pp; English.

XX The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC hGM-CSF fusion protein construct which comprises human PAP  
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2  
 CC signal sequence, mature HER-2 membrane distal extracellular and  
 CC intracellular domains, an Ala Ala linker, a mature human granulocyte-  
 CC macrophage colony stimulating factor (GM-CSF) sequence and a  
 CC C-terminal tag.

XX Sequence 690 AA:

Query Match 100.0%; Score 1587; DB 22; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-131;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELTPLPNASLSFLQDIOEVQ 60  
 DB 41 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELTPLPNASLSFLQDIOEVQ 100  
 QY 61 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVPGASPGGLREIQ 120  
 DB 101 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVPGASPGGLREIQ 160  
 QY 121 LRSLTETLKGVLIQRNPOLCYODTILMKDIFRKNNOALTLTIDTNSRACHCSPCKG 180  
 DB 161 LRSLTETLKGVLIQRNPOLCYODTILMKDIFRKNNOALTLTIDTNSRACHCSPCKG 220  
 QY 181 SRGWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 240  
 DB 221 SRGWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 280  
 QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCATACPNYVLTSDVGS 289  
 DB 281 GICELHCPALVYNTDTFESMPNPEGRTYTGASCATACPNYVLTSDVGS 329

RESULT 10  
 AAE13111

ID AAE13111 standard; protein; 697 AA.

XX AAE13111;

DT 28-JAN-2002 (first entry)

DE Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW Immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;  
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;  
 KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.

XX Chimeric - Homo sapiens.

OS Chimeric - Rattus norvegicus.

XX Chimeric - Unidentified.

XX MO200174855-A2.

XX 11-OCT-2001.  
PD 30-MAR-2001; 2001MO-US10515.  
XX 30-MAR-2000; 2000US-193504P.  
XX 30-MAR-2000; 2000US-193504P.  
XX (DEND-) DENDREON CORP.  
XX Laus R, Vidovic D, Graddis T;  
XX WPI; 2001-662965/76.  
DR N-PSDB; AAD21567.  
XX An immunostimulatory fusion protein comprising the intracellular domain  
PT of HER-2 and an antigen elicits an immune response to the antigen and  
PT is useful for the treatment of associated cancer associated -  
XX Claim 7; Page 27; 59pp; English.  
XX The invention relates to immunostimulatory fusion proteins (IFP) and  
CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
CC polypeptide antigen component and an immunostimulatory component derived  
CC from the intracellular domain of HER-2 protein which is effective to  
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
CC immune response to the antigen. IFP or superactivated dendritic cells  
CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
CC associated with a particularly antigen. The present sequence is HER500  
CC rcm-CSF fusion protein construct which comprises human PAP  
CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2  
CC signal sequence, mature HER-2 membrane distal extracellular domain,  
CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,  
CC HER-2 membrane distal intracellular domain, an Ala Arg linker, a mature  
CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence  
CC and a C-terminal tag.  
XX  
SQ Sequence 697 AA;  
Query Match 100.0%; Score 1587; DB 22; Length 697;  
Best Local Similarity 100.0%; Pred. No. 2.4e-131;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 STOVCTGDMKRLRASPETHLDMRLHLYGCGVQVGNLELTYLPTNASLSFLDIOEVO 60  
DB 41 STOVCTGDMKRLRASPETHLDMRLHLYGCGVQVGNLELTYLPTNASLSFLDIOEVO 100  
QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDGDPPLNNTPTVGASPGGLREIQ 120  
DB 101 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDGDPPLNNTPTVGASPGGLREIQ 160  
QY 121 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180  
DB 161 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 220  
QY 181 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 240  
DB 221 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 280  
QY 241 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 289  
DB 281 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 329  
RESULT 11  
AAB21204  
ID AAB21204 standard; protein: 712 AA.  
XX  
AC AAB21204;  
XX  
DT 12-JAN-2001 (first entry)  
XX  
DE Human HER-2/neu fusion protein.  
XX

KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;  
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
KW colon cancer; fusion protein.  
XX Homo sapiens.  
OS Synthetic.  
PN WO200044899-A1.  
XX  
PD 03-AUG-2000.  
XX  
XX 28-JAN-2000; 2000MO-US02164.  
XX 29-JAN-1999; 99US-0117976.  
XX  
XX (CORI-) CORIXA CORP.  
XX (SMIT) SMITHKLINE BEECHAM.  
XX  
XX Cheever MA, Gheysen D;  
XX WPI; 2000-505976/45.  
DR HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
XX useful for vaccinating against breast, ovarian, colon, lung and  
XX prostate cancers -  
XX Claim 27; Fig 13; 128pp; English.  
XX The present sequence is a fusion protein comprising the extracellular  
XX domain and a preferred portion of the phosphorylation domain of the human  
XX HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase family of  
XX receptor-like glycoproteins and shows homology to the epidermal growth  
XX factor receptor (EGFR). It probably plays a part in cell growth and/or  
XX differentiation. The HER-2/neu gene is an oncogene. HER-2/neu fusion  
XX proteins may be used to treat or prevent cancer by eliciting or enhancing  
XX an immune response to the HER-2/neu protein. They may be used to treat  
XX malignancies such as breast, ovarian, colon, lung and prostate cancers,  
XX and may be used as an antigen to vaccinate against these neoplasias.  
XX  
SQ Sequence 712 AA;  
Query Match 100.0%; Score 1587; DB 21; Length 712;  
Best Local Similarity 100.0%; Pred. No. 2.5e-131;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 STOVCTGDMKRLRASPETHLDMRLHLYGCGVQVGNLELTYLPTNASLSFLDIOEVO 60  
DB 22 STOVCTGDMKRLRASPETHLDMRLHLYGCGVQVGNLELTYLPTNASLSFLDIOEVO 81  
QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDGDPPLNNTPTVGASPGGLREIQ 120  
DB 82 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDGDPPLNNTPTVGASPGGLREIQ 141  
QY 121 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180  
DB 142 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 201  
QY 181 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 240  
DB 202 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 261  
QY 241 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 289  
DB 262 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 310  
RESULT 12  
AAM51149  
ID AAM51149 standard; protein: 712 AA.  
XX  
AC AAM51149;  
XX  
DT 17-JUN-2002 (first entry)  
XX

XX Her-2/neu extracellular domain-delta-phosphorylation domain fusion.  
 DE Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;  
 XX tyrosine kinase; receptor; c-erbB2; gene therapy.  
 KM  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Domain 1..653  
 FT /note="extracellular domain"  
 FT Domain 654..712  
 FT /note="phosphorylation domain fragment"  
 XX  
 PN WO200212341-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 03-AUG-2001: 2001WO-US24283.  
 XX  
 PR 03-AUG-2000: 2000US-0632507.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX (SMK) SMITHKLINE BEECHAM BIOLOGICALS.  
 PI Cheever MA, Gheysen D;  
 XX  
 DR WPI: 2002-241743/29.  
 XX  
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT or enhancing an immune response to the protein, has Her-2/neu  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain -  
 XX  
 PS Claim 37; Fig 13; 141pp; English.  
 XX  
 CC The present sequence is that of a fusion protein between the  
 CC extracellular domain and a fragment (delta) of the phosphorylation  
 CC domain of human Her-2/neu (see AAM5113), an oncogenic self-protein  
 CC and target for anti-cancer vaccines. The fusion protein can be  
 CC obtained by recombinant DNA methods. Her-2/neu overexpression  
 CC correlates with a poor prognosis in breast and ovarian cancers.  
 CC The invention provides Her-2/neu fusion proteins, nucleic acids  
 CC encoding them, viral vectors, and vaccines comprising the fusion  
 CC proteins or nucleic acid molecules. In preferred fusion proteins,  
 CC the extracellular domain of Her-2/neu is fused to a Her-2/neu  
 CC intracellular domain or phosphorylation domain (or its Delta) or  
 CC fragment). An immune response to Her-2/neu protein is elicited or  
 CC enhanced by administering the fusion protein in the form of a vaccine,  
 CC or by transfecting cells of an animal *ex vivo* with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu  
 CC fusion protein can be used to remove tumour cells from a sample in  
 CC order to inhibit the development of cancer in a patient.  
 XX  
 SQ Sequence 712 AA:  
 XX  
 Query Match 100.0%; Score 1587; DB 23; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-131;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 STQVCTGTDMLKRLPASPETHDMLRHLVYOGCCVOGNTLFTLPTNASTSLFDQIDQEVQ 60  
 DB 22 STQVCTGTDMLKRLPASPETHDMLRHLVYOGCCVOGNTLFTLPTNASTSLFDQIDQEVQ 81  
 OY 61 GYVLIAHNOVROVPLQRLRIVRGTQLFEDNYALAVLDNCDPLNTPTVTCASFGGLREIQ 120  
 DB 82 GYVLIAHNOVROVPLQRLRIVRGTQLFEDNYALAVLDNCDPLNTPTVTCASFGGLREIQ 141  
 OY 121 LRSLETLKGVLIQIRNPQLCYODTILMKDIFKNNQALTLIDTNRSRACHCSPCKG 180

DB 142 LRSLETLKGVLIQIRNPQLCYODTILMKDIFKNNQALTLIDTNRSRACHCSPCKG 201  
 OY 181 SRCMGESSEDCOSLTRVYVAGGACGACRGPPLTPOCHQCAAGCTGPKHSCLACLHFNHS 240  
 DB 202 SRCMGESSEDCOSLTRVYVAGGACGACRGPPLTPOCHQCAAGCTGPKHSCLACLHFNHS 261  
 OY 241 GICELHCPALVTYNTDFEESMPNDEGRYTGASCVCAPYNYLSTDVGS 289  
 DB 262 GICELHCPALVTYNTDFEESMPNDEGRYTGASCVCAPYNYLSTDVGS 310  
 RESULT 13  
 AAM19764  
 ID AAM19764 standard; Protein; 782 AA.  
 XX  
 AC AAM19764;  
 XX  
 DT 17-SEP-1997 (first entry)  
 XX  
 DE Her-2-GM-CSF immunostimulant fusion protein.  
 XX  
 KM Her-2-GM-CSF; granulocyte macrophage colony stimulating factor;  
 KM growth factor receptor; oncogene; immunostimulant; cancer;  
 KM therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..653  
 FT Peptide /label= Her2  
 FT /label= 654..655  
 FT Protein /label= Linker  
 FT /label= 656..782  
 FT /label= GM-CSF  
 XX  
 PN MO9724438-A1.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PF 23-DEC-1996: 96WO-US20241.  
 XX  
 PR 28-DEC-1995: 95US-0579823.  
 XX  
 PA (ACTI-) ACTIVATED CELL THERAPY INC.  
 XX  
 PI Laus R, Ruegg CL, Wu H;  
 DR WPI: 1997-363674/33.  
 XX N-PSDB; AAT72725.  
 XX  
 PT Potent APC that activates T-cells to give multivalent cellular  
 PT immune response - can also induce a cytotoxic T-cell response in a  
 PT vertebrate subject  
 XX  
 PS Disclosure; Fig 8; 45pp; English.  
 XX  
 CC A fusion protein (AAM19764) comprises Her2 (a growth factor receptor  
 CC that is overexpressed in breast, ovarian and other cancer cells)  
 CC and granulocyte-macrophage colony stimulating factor (GM-CSF). It  
 CC is the expression product of a nucleic acid molecule (AAT72725)  
 CC prep. by PCR amplification of Her2 cDNA from a breast cancer cell  
 CC line and fusion to GM-CSF cDNA. Fusion expression vectors can be  
 CC used to transfect mammalian and insect cells. The Her2-GM-CSF  
 CC fusion protein is used to generate anti-Her2 immunity. Tumour  
 CC cells are eliminated by cytotoxic T lymphocytes activated *in vivo*  
 CC or *in vitro* by exposure to antigen-presenting cells exposed to the  
 CC fusion protein.  
 XX  
 SQ Sequence 782 AA:  
 XX  
 Query Match 100.0%; Score 1587; DB 18; Length 782;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-131;



Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVGNLLETLYPTNASLSFLDIOEVQ 60  
 |||||||  
 Db 22 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVGNLLETLYPTNASLSFLDIOEVQ 81

QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPVTGASPGGLRELQ 120  
 |||||||  
 Db 82 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPVTGASPGGLRELQ 141

QY 121 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKKG 180  
 |||||||  
 Db 142 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKKG 201

QY 181 SRCWGESSEDCQSLTRTYCAGGCARCKGRLPTDCHEQCAAGCTGPKHSCLACLHFNHS 240  
 |||||||  
 Db 202 SRCWGESSEDCQSLTRTYCAGGCARCKGRLPTDCHEQCAAGCTGPKHSCLACLHFNHS 261

QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289  
 |||||||  
 Db 262 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 14  
 AAB21203  
 ID AAB21203 standard; protein: 919 AA.

AC AAB21203;  
 DT 12-JAN-2001 (first entry)  
 XX  
 DE Human HER-2/neu fusion protein.  
 XX  
 KW Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;  
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KW colon cancer; fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200044899-A1.  
 XX  
 PD 03-AUG-2000.  
 XX  
 PF 28-JAN-2000; 2000WO-US02164.  
 XX  
 PR 29-JAN-1999; 99US-0117976.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM.  
 XX  
 PI Cheever MA, Gheysen D;  
 DR WPI: 2000-505976/45.  
 XX  
 XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
 PT useful for vaccinating against breast, ovarian, colon, lung and  
 PT prostate cancers -  
 XX  
 PS Claim 2; Fig 12; 128pp; English.  
 XX  
 CC The present sequence is a fusion protein comprising the extracellular  
 CC domain and the phosphorylation domain of the human HER-2/neu protein.  
 CC HER-2/neu is a member of the tyrosine kinase family of receptor-like  
 CC glycoproteins and shows homology to the epidermal growth factor receptor  
 CC (EGFR). It probably plays a part in cell growth and/or differentiation.  
 CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used  
 CC to treat or prevent cancer by eliciting or enhancing an immune response  
 CC to the HER-2/neu protein. They may be used to treat malignancies such as  
 CC breast, ovarian, colon, lung and prostate cancers, and may be used as an  
 CC antigen to vaccinate against these neoplasias.  
 XX  
 SQ Sequence 919 AA;

Query Match 100.0%; Score 1587; DB 21; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 3,5e-131;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVGNLLETLYPTNASLSFLDIOEVQ 60  
 |||||||  
 Db 22 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVGNLLETLYPTNASLSFLDIOEVQ 81

QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPVTGASPGGLRELQ 120  
 |||||||  
 Db 82 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPVTGASPGGLRELQ 141

QY 121 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKKG 180  
 |||||||  
 Db 142 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKKG 201

QY 181 SRCWGESSEDCQSLTRTYCAGGCARCKGRLPTDCHEQCAAGCTGPKHSCLACLHFNHS 240  
 |||||||  
 Db 202 SRCWGESSEDCQSLTRTYCAGGCARCKGRLPTDCHEQCAAGCTGPKHSCLACLHFNHS 261

QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289  
 |||||||  
 Db 262 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 15  
 AAM51148  
 ID AAM51148 standard; protein: 919 AA.

AC AAM51148;  
 DT 17-JUN-2002 (first entry)  
 XX  
 DE Her-2/neu extracellular domain-phosphorylation domain fusion.  
 XX  
 KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;  
 KW tyrosine kinase; receptor; c-erbB2; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200212341-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 03-AUG-2001; 2001WO-US24283.  
 XX  
 PR 03-AUG-2000; 2000US-0632507.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Cheever MA, Gheysen D;  
 DR WPI: 2002-241743/29.  
 XX  
 XX Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain -  
 XX  
 PS Claim 2; Fig 12; 141pp; English.  
 XX  
 CC The present sequence is that of a fusion protein between the  
 CC extracellular domain and phosphorylation domain of human Her-2/neu  
 CC (see AAM51143), an oncogenic self-protein and target for anti-cancer  
 CC vaccines. The fusion protein can be obtained by recombinant DNA

CC methods. Her-2/neu overexpression correlates with a poor prognosis  
CC in breast and ovarian cancers. The invention provides Her-2/neu  
CC fusion proteins, nucleic acids encoding them, viral vectors, and  
CC vaccines comprising the fusion proteins or nucleic acid molecules.  
CC In preferred fusion proteins, the extracellular domain of a  
CC Her-2/neu protein is fused to a Her-2/neu intracellular domain or  
CC phosphorylation domain (or its DeltapD fragment). An immune  
CC response to Her-2/neu protein is elicited or enhanced by  
CC administering the fusion protein in the form of a vaccine, or by  
CC transfecting cells of an animal ex vivo with a nucleic acid  
CC encoding the fusion protein, and delivering the transfected cells  
CC to the animal. The fusion proteins, nucleic acids, and isolated  
CC specific T-cells are useful for inhibiting the development of a  
CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
CC in a patient. T cells that specifically react with a Her-2/neu  
CC fusion protein can be used to remove tumour cells from a sample in  
CC order to inhibit the development of cancer in a patient.  
XX

SQ Sequence 919 AA:

Query Match

Best Local Similarity 100.0%; Score 1587; DB 23; Length 919;

Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 STGYCTGTDKRLRPASPEHLDMLRLHYGCGVVGNGLELTYLPTNASLFLDIDIEVQ 60
   |||||||
DB 22 STGYCTGTDKRLRPASPEHLDMLRLHYGCGVVGNGLELTYLPTNASLFLDIDIEVQ 81
   |||||||
QY 61 GYVLIAHNOVROYPLQRLRIYRGSTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 120
   |||||||
DB 82 GYVLIAHNOVROYPLQRLRIYRGSTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 141
   |||||||
QY 121 LRSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNQLALTLDITNRSRACHPCSPMCKG 180
   |||||||
DB 142 LRSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNQLALTLDITNRSRACHPCSPMCKG 201
   |||||||
QY 181 SRMGESSEDCOSLTRVYCGAGCARGPLPTQCCHEGCAAGCTGPKHSDCLAHFNHS 240
   |||||||
DB 202 SRMGESSEDCOSLTRVYCGAGCARGPLPTQCCHEGCAAGCTGPKHSDCLAHFNHS 261
   |||||||
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 289
   |||||||
DB 262 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310
   |||||||
```

Search completed: April 28, 2003, 13:39:46  
Job time : 26.1973 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:38:09 ; Search time 10.6321 Seconds

(without alignments)  
2613.108 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STQVCTGTGTMKLRLPASPER.....FGASCVTACPYNLTSDVGS 289

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: p1r1:\*  
3: p1r2:\*  
4: p1r3:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	1255	1 A24571	protein-tyrosine k
2	1353.5	85.3	1260	1 TVRNU	epidermal growth f
3	1322	83.3	1254	2 I48161	P-185 precursor
4	739	46.6	527	2 A42032	epidermal growth f
5	739	46.6	1223	1 TVCHLV	epidermal growth f
6	716.5	45.1	1210	1 GOHUB	epidermal growth f
7	716.5	45.1	1210	2 A53183	epidermal growth f
8	716.5	45.1	1308	2 A47253	epidermal growth f
9	708.5	44.6	644	2 A36325	epidermal growth f
10	686	43.2	1342	2 A36223	epidermal growth f
11	652	40.1	1339	2 JC4387	kinase-related tra
12	640.5	41.1	1166	1 S06142	epidermal growth f
13	536	33.8	843	2 A27131	protein-tyrosine k
14	396.5	25.0	1323	2 E88257	epidermal growth f
15	396.5	25.0	1374	2 S70712	protein let-23 (lm
16	390.5	24.6	1369	2 S70713	protein-tyrosine k
17	375.5	23.7	1330	1 GOFPE	epidermal growth f
18	341	21.5	366	2 D45558	epidermal growth f
19	341	21.5	1717	1 A45558	epidermal growth f
20	331	20.9	333	2 B45558	epidermal growth f
21	331	20.9	342	2 C45558	epidermal growth f
22	315.5	19.9	1363	2 T43220	insulin-like growth
23	273.5	17.2	1477	2 T18534	insulin-like growth
24	255	16.1	2101	2 S57245	insulin receptor (
25	255	16.1	2148	1 A56081	insulin receptor (
26	246.5	15.5	1300	2 A36502	insulin receptor-r
27	246.5	15.5	1607	2 T43212	insulin-like growth
28	244.5	15.4	1383	2 A36080	insulin receptor p
29	243.5	15.3	1372	2 A34157	insulin receptor p

30	243.5	15.3	1382	1 INHUR	insulin receptor p
31	232.5	14.7	1367	1 IGHUR1	insulin-like growth
32	232.5	14.7	1371	2 A33837	insulin-like growth
33	231.5	14.6	1268	2 B36502	insulin receptor-r
34	228.5	14.4	340	2 B47417	insulin receptor-r
35	224.5	14.1	183	2 JH0803	tyrosine kinase re
36	220.5	13.9	329	2 A48805	insulin-like growth
37	207	13.0	1390	2 T30346	insulin receptor
38	185	11.7	1846	2 T42047	insulin receptor h
39	132.5	8.3	1299	2 T43251	insulin receptor h
40	125	7.9	1548	2 S34583	serine proteinase
41	121	7.6	1680	2 A43434	furin (EC 3.4.21.7
42	109.5	6.9	91	2 I51179	IGF-I receptor - c
43	108.5	6.8	798	2 B28193	integrin beta-1* c
44	107	6.7	915	1 A48225	subtilisin-like pr
45	106.5	6.7	837	2 S43656	furin (EC 3.4.21.7

#### ALIGNMENTS

##### RESULT 1

A:24571 protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human

N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e

C:Species: Homo sapiens (man)

C>Date: 25-Oct-1987 #sequence\_revision 06-Dec-1996 #text\_change 11-Jun-1999

C:Accession: A24571; A25491; A44188; B44188; I59509; I57622

R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T

Nature 319, 230-234, 1986

A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth

A:Reference number: A24571; MUID:66118663; PMID:3003577

A:Accession: A24571

A:Molecule type: mRNA

A:Residues: 1-1255 <TAM>

A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198

R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.

Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985

A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid

A:Reference number: A25491; MUID:66016729; PMID:2955967

A:Accession: A25491

A:Molecule type: DNA

A:Residues: 740-910 <COU2>

A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989

A:Accession: B44188

A:Molecule type: mRNA

A:Residues: 1-517, 'RALU', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>

A:Cross-references: GB:M11730; NID:g183986

R:King, C.R.; Kraus, M.H.; Aaronson, S.A.

Science 229, 974-976, 1985

A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.

A:Reference number: I59509; MUID:85272597; PMID:2992089

A:Accession: I59509

A:Molecule type: DNA

A:Residues: 832-909 <REX>

A:Cross-references: GB:U29395; NID:g459807; PIDN:AAA35809.1; PID:g459808

R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.

Mol. Cell. Biol. 7, 2597-2601, 1987

A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio

A:Reference number: I57622; MUID:87286898; PMID:3039351

A:Accession: I57622

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:9183983; PIDN:AAA58637.1; PID:9553332  
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
 C:Genetics:  
 A:Gene: GDB:ERBB2; NGL; NEU; HER-2  
 A:Cross-references: GDB:120613; OMIM:164870  
 A:Map position: 17q21.1-17q21.1  
 A:Introns: 25/1; 75/3; 147/1; 883/3  
 A:Note: The list of introns is incomplete  
 C:Function:  
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-1255/Product: protein-tyrosine kinase erbb2 #status predicted <MAT>  
 F:22-653/Domain: extracellular #status predicted <EXT>  
 F:70-104/Domain: EGF receptor extracellular domain repeat <EEL>  
 F:195-605/Domain: EGF receptor extracellular domain repeat <EE2>  
 F:654-675/Domain: transmembrane #status predicted <TM>  
 F:676-1255/Domain: intracellular #status predicted <INT>  
 F:18-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase homology motif  
 F:68/124,187,259,530,571,629/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:753/Active site: Lys #status predicted  
 F:1133,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 100.0%; Score 1587; DB 1; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 2,6e-109;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TQVCTGTDKMLRLPASPEHDLMLRLHYOGCQVVGNGLETTYLPPTNASISFLQDIOEVG 60  
 DB 22 TQVCTGTDKMLRLPASPEHDLMLRLHYOGCQVVGNGLETTYLPPTNASISFLQDIOEVG 81  
 OY 61 GYVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120  
 DB 82 GYVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 141  
 OY 121 LRSLETLKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCKMG 180  
 DB 142 LRSLETLKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCKMG 201  
 OY 181 SRWGESSEDCOSLRTFVACGACGACGKGPLPTCCHEQCAAGCTGPKHSDCLCLHFNHS 240  
 DB 202 SRWGESSEDCOSLRTFVACGACGACGKGPLPTCCHEQCAAGCTGPKHSDCLCLHFNHS 261  
 OY 241 GICELHCPALVTYNTDTFESMPNREGRYTFGASCVTACPYNYLSTDVGS 289  
 DB 262 GICELHCPALVTYNTDTFESMPNREGRYTFGASCVTACPYNYLSTDVGS 310

## RESULT 2

TYRRTU

protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 11-Jun-1999

C:Accession: A24562; A61204

R:Barbmann, C.I.; Hung, M.C.; Weinberg, R.A.

Nature 319, 226-230, 1986

A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.

A:Reference number: A24562; MUID:86118662; PMID:3945311

A:Accession: A24562

A:Molecule type: mRNA

A:Residues: 1-1260 &lt;BAR&gt;

A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746

R:Masuul, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,

Carcinogenesis 12, 1975-1978, 1991

A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no

2-thiazolyl[formamide or N-methyl-N-nitrosourea.

A:Reference number: A61204; MUID:92035293; PMID:1682063

A:Accession: A61204

A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 637-663, 'V', 665-702 <MAS>  
 A:Note: authors translated the codon GCA for residue 25 as Val  
 C:Genetics:  
 A:Gene: neu  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>  
 F:658-680/Domain: transmembrane #status predicted <TM>  
 F:723-988/Domain: protein kinase homology <KIN>  
 F:731-739/Region: protein kinase ATP-binding motif  
 F:11,191,263,535,576,634/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted  
 F:758/Active site: Lys #status predicted  
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 85.3%; Score 1353.5; DB 1; Length 1260;  
 Best Local Similarity 85.3%; Pred. No. 4,4e-92;  
 Matches 247; Conservative 13; Mismatches 28; Indels 1; Gaps 1;

OY 2 TQVCTGTDKMLRLPASPEHDLMLRLHYOGCQVVGNGLETTYLPPTNASISFLQDIOEVG 61  
 DB 26 TQVCTGTDKMLRLPASPEHDLMLRLHYOGCQVVGNGLETTYLPPTNASISFLQDIOEVG 85  
 OY 62 YVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120  
 DB 86 YVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 145  
 OY 121 LRSLETLKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCKMG 180  
 DB 146 LRSLETLKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCKMG 205  
 OY 181 SRWGESSEDCOSLRTFVACGACGACGKGPLPTCCHEQCAAGCTGPKHSDCLCLHFNHS 240  
 DB 206 SRWGESSEDCOSLRTFVACGACGACGKGPLPTCCHEQCAAGCTGPKHSDCLCLHFNHS 265  
 OY 241 GICELHCPALVTYNTDTFESMPNREGRYTFGASCVTACPYNYLSTDVGS 289  
 DB 266 GICELHCPALVTYNTDTFESMPNREGRYTFGASCVTACPYNYLSTDVGS 314

## RESULT 3

I48161

P-185 precursor - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Jun-1999

C:Accession: I48161

R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishika

Gene 140, 251-255, 1994

A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.

A:Reference number: I48161; MUID:94193007; PMID:7908275

A:Accession: I48161

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1254 &lt;RES&gt;

A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595

C:Genetics:

A:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP

F:718-983/Domain: protein kinase homology &lt;KIN&gt;

F:726-734/Region: protein kinase ATP-binding motif

Query Match 83.3%; Score 1322; DB 2; Length 1254;  
 Best Local Similarity 83.3%; Pred. No. 9,1e-90;  
 Matches 240; Conservative 16; Mismatches 32; Indels 0; Gaps 0;

OY 2 TQVCTGTDKMLRLPASPEHDLMLRLHYOGCQVVGNGLETTYLPPTNASISFLQDIOEVG 61  
 DB 23 TQVCTGTDKMLRLPASPEHDLMLRLHYOGCQVVGNGLETTYLPPTNASISFLQDIOEVG 82  
 OY 62 YVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 121

```

Db      83 YMLAHSGVRRVRFQRLRIYVGTQLFEDKALVALDNRPDLNVATTAIGTPEGLELD 142
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      122 RSLTEILKGGVLIQNIQPOLCYODTILMKDIFHKNNQIATLITDITNSRACHPCSPMKGS 181
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      143 RSLTEILKGGVLIQNIQPOLCYODTIVLMKVDYFRKNNQIAPVDITNSRACPPAPACKDN 202
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      182 RCMGSESDCSLIRTYCAGGCAKRCGLPLTDCCHEGCAAGCTGPKHSDCLACIAPHNSG 241
      ||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      203 HCMKASPEDCOTLTGTLAPRAVPAARARLPTDCCHEGCAAGCTGPKHSDCLACIAPHNSG 262
      ||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      242 ICCLHCPALVTYNTDFESMPNPEGRITTFASCAVTLACPYNYLTSDVGS 289
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      263 ICCLHCPALVTYNTDFESMPNPEGRITTFASCAVTLCPYNYLTSDVGS 310
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 4  
A42032  
epidermal growth factor receptor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 31-Dec-1993 #sequence:revision 31-Dec-1993 #text:change 18-Jun-1999  
C:Accession: A42032  
R:Flickinger, T.W.; Mainle, N.J.; Kung, H.J.  
Mol. Cell. Biol. 12, 883-893, 1992  
A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, t

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-527 <FL>  
A:Cross-references: GB:1767637; NID:9211737; PIDN:AAA48759.1; PID:9211738  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBI:76892, NCBI:76893)  
A:Superfamily: epidermal growth factor receptor; protein kinase homology  
Keywords: AMP; growth factor receptor

Query Match	46.68;	Score 739;	DB 2;	Length 527;
Best Local Similarity	48.18;	Pred. No. 3.9e-47;		
Matches 139; Conservative	48;	Mismatches 90;	Indels 12;	Gaps 5;

QY	3	QVCGTGMKLLPASPENHDMLEKOGGVOVNNLETLFYLPTNMSLSTDDIOEYQY	62
QY	38	KVCGSTNNKLTQLGHVEDHFLSLDQRMATNKEEYVLSNLETLVEHNRDLPTLKTIOEYAGY	97
Db	63	VLIAHNOVROYPLQRLRIYRGTOLEFEDNYALAVLDNGDPNLTNTPYTGASPGELREIQJR	122
QY	98	VLIAMNVDYIPRENLDQIRGNVLDSPALVLSNTH - NKKIQ - - - - - GLRELPMK	148
Db	123	SLTELKGGVLIQNRPOLCOTDTLTKMDIFHHKNNQLALTLID - TNSRACHPCSPMKGS	181
QY	149	RLESLINGGVYSINNNPKLCNNMDYVLAMNDIIDTTSRK - PLTVLDEASNLSSCPKCHPNCTED	207
Db	182	RCMGESSDDQSILRTYVAGGCA - RCKGRPLTPOCHGCGCAAGCTGPRKHSCLACLFPHNS	240
QY	208	HCMAGAEQNCCTLTKYVITCAQGSRCRKGKVPASDCCHNQAAGCTGPREPSCLACRFRFD	267

Oy 241 GICELHPALVYNTDTEFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 289  
| : | | | | : : | | | : | | | | : | | |  
Db 268 ATCKDTCPLVLYNPTTYQMDVNPEGKYSFGATCVRECPHNYVVTDHGS 316

RESULT 5  
TVCMLV  
external growth factor receptor precursor - chicken  
N:contams: protein-tyrosine kinase (EC 2.7.1.11) erbb  
C:Species: Gallus gallus (chicken)  
C:Date: 28-Feb-1986 #sequence\_revision 05-May-1995 #text\_change 04-Feb-2000  
C:Accession: A27720: A00643

MOL. CELL. BIOL. 8, 1970-1978, 1988  
A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mouse cells  
A:Reference number: A27720; MUID:88261272, PMID:3260329  
A:Accession: A27720

A:Molecule type: mRNA  
A:Residues: 1-1223 <LAX>  
A:Cross-references: GB:M20386  
R:Nilssen, T.W.; Maroney, P.A.; Goodwin, R.G.; Roltman, F.M.; Riltenden, L.B.; Raines  
Cell 41, 719-726, 1985  
A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and  
A:Reference number: A00643; MUID:85228222, PMID:2968784  
A:Accession: A00643  
A:Molecule type: mRNA  
A:Residues: 585-1223 <NIL>  
A:Cross-references: GB:M10066

C:Superfamily:epidermal growth factor receptor; protein kinase homology  
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor  
specific protein kinase  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-1223/Product: epidermal growth factor receptor #status predicted <MAT>  
F:31-654/Domain: extracellular #status predicted <EXT>  
F:81-307/Domain: EGF receptor extracellular domain repeat <EE1>  
F:397-610/Domain: EGF receptor extracellular domain repeat <EE2>  
F:655-677/Domain: transmembrane #status predicted <TM>  
F:678-1223/Domain: intracellular #status predicted <INT>  
F:719-984/Domain: protein kinase homology <KIN>  
F:727-735/Region: protein kinase ATP-binding motif  
F:136-202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #status predicted  
F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
F:75/Active site: Lys #status predicted  
F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #

Query Match	46.6%	Score 739	DB 1	length 1223
Best Local Similarity	48.1%	Pred. No.	9e-47	
Matches 139	Conservative 48	Mismatches 90	Indels 12	Gaps 5

[illegible]

RESULT 6  
GOHUE  
epidermal growth factor receptor precursor - human

Citation: 15-Mo1984 [sequence revision 27-Nov-1985 #text change 11-Jun-1999 .  
Accession: A00641: A25772: S30024: A38672: A00642: A43615: A23062: A05281: A60143:  
R.Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.  
rg, F.H.  
Nature 309, 418-425, 1984  
A>Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression  
A.Reference number: A00641; MUID:84219729;PMID:6328312

A: Molecule type: mRNA  
A: Residues: 1-1210 <URL>  
A: Cross-references: EMBL: X00568; NID: g31113; PDB: g75792  
A: Note: the authors translated the codon AAG for residue 540 as Asn

R: Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985  
 A:Title: Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene  
 A:Reference number: A25772; PMID:85270438; PMID:2991899  
 A:Accession: A25772  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-29 <ISH>  
 A:Cross-references: GB:M1234; NID:q181981; PIDN:AAA52370.1; PID:g553272  
 R: Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ullrich, A.; Waterfield, M.  
 Oncogene Res. 1, 375-396, 1987  
 A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification of the coding sequence  
 A:Reference number: S30024; PMID:88217353; PMID:3329716  
 A:Accession: S30024  
 A:Molecule type: DNA  
 A:Residues: 1-29 <HA2>  
 A:Cross-references: EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID:g31119  
 R: Haley, J.D.; Waterfield, M.D.  
 J. Biol. Chem. 266, 1746-1753, 1991  
 A:Title: Contributory effects of de Novo transcription and premature transcript termination on the expression of the human EGF receptor gene  
 A:Reference number: A38672; PMID:91107677; PMID:1988448  
 A:Accession: A38672  
 A:Molecule type: DNA  
 A:Residues: 1-29 <HNL>  
 A:Cross-references: GB:M38425; NID:q181977; PIDN:AAA63171.1; PID:g553271  
 A:Experimental source: carcinoma cell line A431-7  
 R: Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Merlino, I.  
 Nature 309, 806-810, 1984  
 A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs  
 A:Reference number: A00642; PMID:84245855; PMID:6330563  
 A:Accession: A00642  
 A:Molecule type: mRNA  
 A:Residues: 150-187, 'KSVIGAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-321, '798-799', 'TP', 802-811, 'R', 813-942 <XUY>  
 A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF receptor  
 R: Lin, C.R.; Chen, W.S.; Krulger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verna, I.  
 Science 224, 843-848, 1984  
 A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplification  
 A:Reference number: A43615; PMID:84196372; PMID:6326261  
 A:Accession: A43615  
 A:Molecule type: mRNA  
 A:Residues: 713-964 <LIN>  
 A:Experimental source: epidermoid carcinoma cell line A431  
 R: Stamen, F.A.; Gope, M.L.; Schultz, T.2.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
 Biochem. Biophys. Res. Commun. 124, 125-132, 1984  
 A:Reference number: A23062; PMID:85046483; PMID:6093780  
 A:Accession: A23062  
 A:Molecule type: mRNA  
 A:Residues: 1028-1210 <SIM>  
 R: Weber, W.; Gill, G.N.; Speiss, J.  
 Science 224, 294-297, 1984  
 A:Reference number: A05281; PMID:84172183; PMID:6324343  
 A:Accession: A05281  
 A:Molecule type: protein  
 A:Residues: 25-30 'S', 32-51, 454-467 <WEB>  
 R: Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.  
 J. Biol. Chem. 260, 5205-5208, 1985  
 A:Title: Identification of residues in the nucleotide binding site of the epidermal growth factor receptor  
 A:Reference number: A60143; PMID:85182650; PMID:2985580  
 A:Accession: A60143  
 A:Molecule type: protein  
 A:Residues: 740-744, 'X', 746-747 <RUS>  
 R: Roczniowski, B.; Mosig, G.; Cohen, S.  
 Nature 309, 270-273, 1984  
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superoxide anion  
 A:Reference number: A38023; PMID:84191554; PMID:6325348  
 A:Accession: A38023  
 A:Molecule type: protein  
 A:Residues: 1-714 <RAV>  
 R: Chen, W.S.; Lazear, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.  
 Cell 59, 33-43, 1989  
 A:Title: Functional independence of the epidermal growth factor receptor from a domain involved in internalization  
 A:Reference number: A3331; PMID:90003233; PMID:2790960  
 A:Accession: A3331  
 A:Molecule type: protein  
 A:Residues: 1-714 <RAV>  
 R: Chen, W.S.; Lazear, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.  
 Cell 59, 33-43, 1989  
 A:Title: Functional independence of the epidermal growth factor receptor from a domain involved in internalization  
 A:Reference number: A3331; PMID:90003233; PMID:2790960  
 A:Accession: A3331  
 A:Molecule type: protein

C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor  
 C:Genetics:  
 A:Gene: GDB:EGFR  
 A:Cross-references: GDB:120610; OMIM:131550  
 A:Map position: 7p12.3-7p12.1  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-1210/Product: EGF receptor #status predicted <EMT>  
 F:25-645/Domain: extracellular #status predicted <EXT>  
 F:390-600/Domain: EGF receptor extracellular domain repeat <EE1>  
 F:646-668/Domain: EGF receptor extracellular domain repeat <EE2>  
 F:646-668/Domain: transmembrane #status predicted <TM>  
 F:669-1210/Domain: intracellular #status predicted <INT>  
 F:710-975/Domain: protein kinase homology <KIN>  
 F:718-726/Region: protein kinase ATP-binding motif  
 F:999-1046/Region: coated-pit mediated internalization signal  
 F:1047-1210/Region: inhibitory  
 F:128-175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status pre  
 F:745/Active site: Lys #status experimental  
 Query Match 45.1%; Score 716.5; DB 1; Length 1210;  
 Best Local Similarity 46.5%; Pred. No. 4, 1e-45;  
 Matches 134; Conservative 44; Mismatches 99; Indels 11; Gaps 2;  
 QY 3 QVGTGDMKRLPASPETHLDMRLHYGGVGGVGNLETLPTNLSLFLQDIQEVGY 62  
 DB 29 KVGQSGNKLQGTGFEDHFLSLQRMFNCEVVLGNLEIYVGRVNDLSPLKTIQEVAGY 88  
 QY 63 VILAHNRYRVPYQRLRIYVGTQLFEDNVALVLDNDPLNNTPTVYGAPEGRLQLR 122  
 DB 89 VILALNVERIPLENDQITGNWYENSVALVLSND-----AKRTGLKELPMR 138  
 QY 123 SLTEILKGVGLIQRNPOLCYQDTILMKDIFPKNNQLATLIDTRSRACPCSPMKGR 182  
 DB 139 NIDELHGNARFNNRNLCAVESIQMDIYSSDFLSMNSMDPQNHLSGCKCPSCNGS 198  
 QY 183 CGEBSSEDCOSLTRFYCAGCA-RCKGRLPTDCCHQCAAGCTGPKHSDCLAEHNSG 241  
 DB 199 CMGAGEENCKKLTICACQCCSCRCRSPSCDCHNCAAGCTGPRSPDCTVCRKFEDEA 258  
 QY 242 ICELHCPALVTYNDPESMPNPEGRRTFGACVTCAPYVLSYDSGS 289  
 DB 259 TCKDTCPPLMLNPTTYQMDVNEBGRKTSFGATCYKCKPRNYVTYDHS 306  
 RESULT 7  
 A53183  
 epidermal growth factor receptor precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999  
 C:Accession: A53183; A43818; S24942; S45325; I49643  
 R: Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.  
 Genes Dev. 8, 399-413, 1994  
 A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor  
 A:Reference number: A53183; PMID:94170966; PMID:8125255  
 A:Accession: A53183  
 A:Molecule type: mRNA  
 A:Residues: 1-1210 <LUE>  
 A:Cross-references: GB:U03425  
 R: Raviv, A.; Lay, I.; Ullrich, A.; Schlesinger, J.; Givol, D.; Morse, B.  
 Oncogene 6, 673-676, 1991  
 A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding  
 A:Reference number: A43818; PMID:91323866; PMID:2030916  
 A:Accession: A43818  
 A:Molecule type: mRNA  
 A:Residues: 1-714 <RAV>  
 A:Cross-references: GB:X59698  
 R: Bislinger, D.P.; Serrero, G.  
 submitted to the EMBL Data Library, June 1992  
 A:Reference number: S24942  
 A:Accession: S24942  
 A:Molecule type: mRNA









A: Experimental source: strain N2  
 R: Koga, M.  
 submitted to the EMBL Data Library, July 1995  
 A: Reference number: S73101  
 A: Accession: S73101  
 A: Molecule type: DNA  
 A: Residues: 1-50, 'G', 52-1374 <KOG>  
 A: Cross-references: EMBL:D63426; NID:g1407562; PIDN:BAA09729.1; PID:g1407563  
 A: Experimental source: strain N2  
 R: Arolan, R.V.; Koga, M.; Mendel, J.E.; Ohshima, Y.; Sternberg, P.W.  
 Nature 348, 693-699, 1990  
 A: Title: The let-23 gene necessary for *Caenorhabditis elegans* vulval induction encodes a  
 A: Reference number: S13422; MUID:91080919; PMID:1979659  
 A: Accession: S13422  
 A: Molecule type: mRNA  
 A: Residues: 52-1374 <ARO>  
 R: Thomas, K.  
 submitted to the EMBL Data Library, March 1996  
 A: Reference number: Z20404  
 A: Accession: T27682  
 A: Status: preliminary; translated from GB/EMBL/DBJ  
 A: Molecule type: DNA  
 A: Residues: 52-1374 <NTL>  
 A: Cross-references: EMBL:Z70038; PIDN:CAA93882.1; GSPDB:GN00020; CESP:ZK1067.1  
 A: Experimental source: clone ZK1067  
 C: Genetics:  
 A: Gene: let-23; CESP:ZK1067.1  
 A: Map position: 2  
 A: Introns: 44/1; 51/1; 90/1; 112/3; 165/3; 217/1; 290/1; 379/1; 418/1; 448/2; 565/1; 608  
 C: Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 C: Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki  
 F: 1-28/Domain: signal sequence #status predicted <SIG>  
 F: 29-1374/Product: protein-tyrosine kinase let-23 #status predicted <MAT>  
 F: 934-1199/Domain: protein kinase homology <KIN>  
 F: 942-950/Region: protein kinase ATP-binding motif

Query Match 25.0%; Score 396.5; DB 2; Length 1374;  
 Best Local Similarity 28.8%; Pred. No. 1.9e-21;  
 Matches 89; Conservative 54; Mismatches 113; Indels 53; Gaps 9;

```

QY 4 VCTGTDMKRLPASPETHLDMRLHLYQGCVVGNLELYLPN----- 47
DB 90 LSGGTNGISRYGTGNI-LEDETMYRGCRVYGNLEITWIEANEIKKMRSTNSTVDPR 148
QY 48 -----ASLSFLDDIOEVQGYVLIANQVRYQVPLQRLRIVRGTFQDFEDNYALAVLDNGDP 101
DB 149 NEDSPLKSIINFEDNLEIRGSLIYRANIOKISFPLRVLYGDEVEFDN-ALYIHKNDK- 206
QY 102 LNNTPVTGASPGGLRELQRLSTELLKGVLIQRRNPOLCY-ODTILMKDIFHKNNQAL 160
DB 207 -----VHEVVMRELVRIRNGSVTTIQDNPKNICYIGDKIDMKELLYDPD--VQ 250
QY 161 TLIDTNRSRACH-----PCSPMKSGSRQWGESSEDCOSLTFRTVCAGGCARC---KGPL 210
DB 251 KVTETNSHQHCYQNGSKMAKCHESC-NDKQWGSQNDQCRVYRSVCPKSGSQCFYSNSTS 309
QY 211 PTDCGHEGCAAGCTGPKHSDCLACLFHNHSGICELHCPALVYNTDTFESMPNPEGRTYF 270
DB 310 SYECCDSACIGGCTGHPNCIACSKYEIDGICIEFCPSRKIFNHKTGRLVFNPDRYQN 369
QY 271 GASCVTACP 279
DB 370 GNHCVKECP 378
  
```

Search completed: April 28, 2003, 13:42:29  
 Job time : 17.6321 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:23 ; Search time 4.53595 Seconds  
(without alignments)  
1984.228 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182  
Sequence: 1 GAGGMVHNRHRSSTRSGG.....STFKGPTAENPEYLGLDVP 217

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1177	99.6	1255	1	ERRB2_HUMAN
2	966.5	81.8	1254	1	ERRB2_MESAU
3	954	80.7	1257	1	ERRB2_RAT
4	283	23.9	245	1	ERRB2_MOUSE
5	164	13.9	1210	1	EGFR_HUMAN
6	157	13.3	1308	1	ERRB4_HUMAN
7	152	12.9	1308	1	ERRB4_RAT
8	145.5	12.3	707	1	SEPO_HUMAN
9	138	11.7	634	1	ERBBB_ALV
10	135.5	11.5	331	1	PRP1_HUMAN
11	135.5	11.5	553	1	ODO2_MYCTU
12	135	11.4	620	1	EXTN_TOBAC
13	134	11.3	1210	1	EGFR_MOUSE
14	130.5	11.0	886	1	SM6B_MOUSE
15	129.5	11.0	251	1	PRP2_HUMAN
16	128.5	10.9	1321	1	IRS2_MOUSE
17	127.5	10.8	3149	1	TEGU_EBV
18	125.5	10.6	2805	1	MAPA_HUMAN
19	124.5	10.5	775	1	ICP0_HSV1
20	124	10.5	503	1	WAP1_HUMAN
21	124	10.5	1106	1	GLI1_HUMAN
22	122	10.3	497	1	MAS2_HUMAN
23	122	10.3	518	1	TPM4_DROME
24	122	10.3	656	1	DNAA_STRO
25	121.5	10.3	938	1	ENB4_EBV
26	120.5	10.2	449	1	APC_BRANA
27	120	10.2	296	1	PRP3_MOUSE
28	120	10.2	426	1	EXLP_TOBAC
29	120	10.2	699	1	VGIG_HSV2
30	119.5	10.1	296	1	CCO1_CAEEL
31	119.5	10.1	1219	1	YW29_YEAST
32	119	10.1	279	1	Y091_NPVP
33	119	10.1	439	1	XP2_XENLA

34	119	10.1	960	1	FGD1_MOUSE
35	118.5	10.0	604	1	RCO1_NEUR
36	118	10.0	261	1	PRP2_MOUSE
37	118	10.0	419	1	SK11_MOUSE
38	118	10.0	2167	1	SHK1_RAT
39	117.5	9.9	276	1	PRPL_HUMAN
40	117	9.9	1324	1	IRS2_HUMAN
41	117	9.9	1664	1	SLP1_CLOTM
42	116.5	9.9	561	1	ERS_HUMAN
43	116	9.8	1248	1	DIAL_HUMAN
44	116	9.8	1274	1	ENAM_MOUSE
45	115.5	9.8	560	1	ERS_MOUSE

## ALIGNMENTS

RESULT 1  
ERRB2\_HUMAN  
ID ERRB2\_HUMAN STANDARD: PRT; 1255 AA.  
AC P04626;  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell  
DE surface receptor HER2) (MLN 19).  
CN ERBB2 OR HER2 OR NGL OR NEU.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-86118663; PubMed-3003577;  
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,  
RA Saito T., Toyoshima K.,  
RT "Similarity of protein encoded by the human c-erbB-2 gene to  
RT epidermal growth factor receptor";  
RL Nature 319:230-234(1986).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-86016729; PubMed-2999974;  
RA Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,  
RA McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,  
RA Francke U., Levinson A., Ullrich A.,  
RT "Tyrosine kinase receptor with extensive homology to EGF receptor  
RT shares chromosomal location with neu oncogene";  
RL Science 230:1132-1139(1985).  
[3]  
RP SEQUENCE OF 737-1031 FROM N.A.  
RX MEDLINE-86016729; PubMed-2999974;  
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.,  
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the  
RT c-erbB-1/epidermal growth factor receptor gene and is amplified in a  
RT human salivary gland adenocarcinoma";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).  
[4]  
RP VARIANTS VAL-654 AND VAL-655.  
RX MEDLINE-93194196; PubMed-8095488;  
RA Ehsani A., Low J., Wallace R.B., Wu A.M.,  
RT "Characterization of a new allele of the human ERBB2 gene by allele-  
RT specific competition hybridization";  
RL Genomics 15:426-429(1993).  
-i- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,  
ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
ALPHA AND AMPHIREGULIN.  
-i- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
tyrosine phosphate.  
-i- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
(POTENTIAL).  
-i- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- P-TM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN  
 CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY  
 CC OF 0.783; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;  
 CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M11767; AAA35808.1; -.  
 DR EMBL: M11761; AAA35808.1; JOINED.  
 DR EMBL: M11762; AAA35808.1; JOINED.  
 DR EMBL: M11763; AAA35808.1; JOINED.  
 DR EMBL: M11764; AAA35808.1; JOINED.  
 DR EMBL: M11765; AAA35808.1; JOINED.  
 DR EMBL: M11766; AAA35808.1; JOINED.  
 DR EMBL: M11730; AAA75493.1; -.  
 DR EMBL: M12036; AAA35978.1; -.  
 DR EMBL: X03363; CAA27060.1; -.  
 DR PIR: A25491; A25491.  
 DR PIR: A24571; A24571.  
 DR HSSP: P11362; 1FGK.  
 DR Genew: HGNC:3430; ERBB2.  
 DR MIM: 164870; -.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Furin-1ike.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR Pfam: PF00757; Furin-1like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR Pfam: PF02757; YLP; 2.  
 DR ProDom: PD000001; Euk\_Pkinase; 1.  
 DR SMART: SM00261; FU; 3.  
 DR SMART: SM00219; TYRKC; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transmembrane; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Polymorphism.  
 KM STGNL 1 21  
 FT CHAIN 22 1255  
 FT DOMAIN 22 652  
 FT TRANSMEM 653 675  
 FT DOMAIN 676 1255  
 FT DOMAIN 720 987  
 FT NP\_BIND 726 734  
 FT BINDING 753 753  
 FT ACT\_SITE 845 845  
 FT DISULFID 195 204  
 FT DISULFID 199 212  
 FT DISULFID 220 227  
 FT DISULFID 224 235  
 FT DISULFID 236 244  
 FT DISULFID 240 252  
 FT DISULFID 255 264  
 FT DISULFID 268 295  
 FT DISULFID 299 311  
 FT DISULFID 315 331  
 FT DISULFID 334 338  
 FT DISULFID 511 520  
 FT DISULFID 515 528  
 FT DISULFID 531 540  
 FT DISULFID 544 560  
 FT DISULFID 560 560

FT DISULFID 563 576  
 FT DISULFID 567 584  
 FT DISULFID 587 596  
 FT DISULFID 600 623  
 FT DISULFID 606 634  
 FT DISULFID 630 642  
 FT MOD\_RES 1139 1139  
 FT MOD\_RES 1248 1248  
 FT CARBOHYD 68 68  
 FT CARBOHYD 124 124  
 FT CARBOHYD 187 187  
 FT CARBOHYD 259 259  
 FT CARBOHYD 530 530  
 FT CARBOHYD 571 571  
 FT CARBOHYD 629 629  
 FT VARIANT 654 654  
 FT VARIANT 655 655  
 FT VARIANT 655 655  
 FT CONFLICT 1170 1170  
 FT CONFLICT 1170 1170  
 FT SEQUENCE 1255 AA; 137909 MM; 39E9DEDA04DC9662 CRC64;  
 SO SEQUENCE 1255 AA; 137909 MM; 39E9DEDA04DC9662 CRC64;  
 Query Match 99.6%; Score 1177; DB 1; Length 1255;  
 Best Local Similarity 99.5%; Pred. No. 3.8e-64;  
 Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GAGGMVHRRSSSTRSGGDLTLGLPSESEAPRSLPASEGAGSDVPDGLGMAK 60  
 DB 1038 GAGGMVHRRSSSTRSGGDLTLGLPSESEAPRSLPASEGAGSDVPDGLGMAK 1097  
 QY 61 LQSLPTHPSPDLQYSDPTVPPLPSETDGYVAPLTCSPOEYVNPQVPYRQPPSPRGPL 120  
 DB 1098 LQSLPTHPSPDLQYSDPTVPPLPSETDGYVAPLTCSPOEYVNPQVPYRQPPSPRGPL 1197  
 QY 121 PAARPAATLEPRKATLSPKNGVYKDVFAFGAENVEYLTPOGAPAPHPAPSPAF 180  
 DB 1158 PAARPAATLEPRKATLSPKNGVYKDVFAFGAENVEYLTPOGAPAPHPAPSPAF 1217  
 QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
 DB 1218 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 1254  
 RESULT 2  
 ID ERB2\_MESAU STANDARD; PRT; 1254 AA.  
 AC 060553;  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (neu proto-oncogene) (C-erbB-2).  
 GN ERBB2 OR NEU.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OC NCBI\_Taxid-10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Nerve;  
 RX MEDLINE-94193007; PubMed-7908275;  
 RA Nakamura T., Ushijima T., Ishizaka Y., Naga M., Arai M.,  
 RA Yamazaki Y., Ishikawa T.;  
 RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";  
 RL Gene 140:251-255(1994).  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,  
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS

(POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D16295; BAA03801.1; -.  
 DR HSSP: P11362; IRGK.  
 DR InterPro: IPR000494; EGFRL\_domain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; Pkinase\_1.  
 DR Pfam: PF00757; Furin-like\_1.  
 DR Pfam: PF01030; Recep\_L\_domain\_2.  
 DR Pfam: PF02757; YLP\_2.  
 DR PRODOM: PD000001; Euk\_Pkinase; 1.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP\_1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR\_1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM\_1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW proto-oncogene; Disease mutation.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 1254 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.  
 FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).  
 FT TRANSEM 653 675 POTENTIAL.  
 FT DOMAIN 676 1254 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 158 368 CYS-RICH.  
 FT DOMAIN 472 644 CYS-RICH.  
 FT DOMAIN 720 987 PROTEIN KINASE.  
 FT NP\_BIND 726 734 ATP (BY SIMILARITY).  
 FT BINDING 753 753 ATP (BY SIMILARITY).  
 FT ACI\_SITE 845 845 BY SIMILARITY.  
 FT DISULFID 195 204 BY SIMILARITY.  
 FT DISULFID 199 212 BY SIMILARITY.  
 FT DISULFID 236 244 BY SIMILARITY.  
 FT DISULFID 240 252 BY SIMILARITY.  
 FT DISULFID 255 264 BY SIMILARITY.  
 FT DISULFID 268 295 BY SIMILARITY.  
 FT DISULFID 299 311 BY SIMILARITY.  
 FT DISULFID 315 331 BY SIMILARITY.  
 FT DISULFID 334 338 BY SIMILARITY.  
 FT DISULFID 511 520 BY SIMILARITY.  
 FT DISULFID 515 528 BY SIMILARITY.  
 FT DISULFID 531 540 BY SIMILARITY.  
 FT DISULFID 544 560 BY SIMILARITY.  
 FT DISULFID 563 576 BY SIMILARITY.  
 FT DISULFID 567 584 BY SIMILARITY.  
 FT DISULFID 587 596 BY SIMILARITY.  
 FT DISULFID 600 623 BY SIMILARITY.  
 FT DISULFID 626 634 BY SIMILARITY.  
 FT DISULFID 630 642 BY SIMILARITY.  
 FT MOD\_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 125 125 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 530 530 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 571 571 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 629 629 N-LINKED (GLCNAC...) (POTENTIAL).

FT VARIANT 658 658 V -> E (IN ONCOGENIC NEU).  
 FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).  
 SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2B81 CRC64;  
 Query Match 81.8%; Score 966.5; DB 1; Length 1254;  
 Best Local Similarity 82.0%; Pred. No. 1.86-51;  
 Matches 178; Conservative 13; Mismatches 25; Indels 1; Gaps 1;  
 QY 1 GAGGWHHRHSSSRSGGDLTGLPSEEPAPSPPLASPGAGSDVFDGLGAAKG 60  
 DB 1038 GAGTAHRRHSSSTRSGGDLTGLPSEEPAPSPPLASPGAGSDVFDGLGAAKG 1097  
 QY 61 LQSLFTDPSPLQRYSEDPYPLPSETDGYVAPLTCSPQPEVYNQDVPPPPSPREGPL 120  
 DB 1098 POSTSPRLSPQLQRYSEDPYPLPSETDGYVAPLTCSPQPEVYNQDVPPPPSPREGPL 1157  
 QY 121 PAAPAGATLEBAKTLSTGKNGVYKADVAFGAVENPEYLPQGCAPQPPAPFSPAF 180  
 DB 1158 PPVRPAGATLEBAKTLSTGKNGVYKADVAFGAVENPEYLPQGCAPQPPAPFSPAF 1216  
 QY 181 DNLVYWDODPPERGAPSTFTGTPAENPEYLGIDVP 217  
 DB 1217 DNLVYWDODPPERGAPSTFTGTPAENPEYLGIDVP 1233  
 RESULT 3  
 ERBB2\_RAT  
 ID ERBB2\_RAT STANDARD: PRT: 1257 AA.  
 AC P06494;  
 DT 01-JAN-1998 (Rel. 06, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor  
 DE receptor-related protein).  
 GN ERBB2 OR NEU.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Neuroblastoma;  
 RX MEDLINE=86118662; PubMed=3945311;  
 RA Bargmann C.I., Hung M.-C., Weinberg R.A.;  
 RT "The neu oncogene encodes an epidermal growth factor receptor-related  
 RT protein.";  
 RL Nature 319:226-230(1986).  
 RN [2]  
 RP SEQUENCE OF 852-905 FROM N.A.  
 RC TISSUE=Sciatic nerve;  
 RX MEDLINE=91222560; PubMed=2025425;  
 RA Lal C., Lemke G.;  
 RT "An extended family of protein-tyrosine kinase genes differentially  
 RT expressed in the vertebrate nervous system.";  
 RL Neuron 6:691-704(1991).  
 RN [3]  
 RP STRUCTURE BY NMR OF 650-668  
 RX MEDLINE=92155181; PubMed=1346763;  
 RA Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,  
 RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;  
 RT "Three dimensional structure of the transmembrane region of the proto-  
 RT oncogenic and oncogenic forms of the neu protein.";  
 RL EMBO J. 11:43-48(1992).  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX.  
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
 CC ALPHA AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS.  
 CC THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- P1M: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X03362; CAA27039.1; ALT\_INIT.  
 DR PIR: A24562; TVRTNU.  
 DR HSSP: P11362; 1FGK.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR Pfam: PF02757; YLP; 2.  
 DR ProDom: PD000001; Euk\_kinase; 1.  
 DR SMART: SM00261; F0; 3.  
 DR SMART: SM00219; TyrKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ARP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KM Transmembrane: Glycoprotein: Multigene family; Receptor; Signal;  
 KM transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KM Proto-oncogene; Disease mutation.  
 FT SIGNAL 1 21  
 FT CHAIN 22 1257  
 FT DOMAIN 22 654  
 FT TRANSMEM 655 677  
 FT DOMAIN 678 1257  
 FT DOMAIN 159 369  
 FT DOMAIN 473 646  
 FT DOMAIN 722 989  
 FT NP\_BIND 728 736  
 FT BINDING 755 755  
 FT ACT\_SITE 847 847  
 FT DISULFID 196 205  
 FT DISULFID 200 213  
 FT DISULFID 221 228  
 FT DISULFID 225 236  
 FT DISULFID 237 245  
 FT DISULFID 241 253  
 FT DISULFID 256 263  
 FT DISULFID 269 296  
 FT DISULFID 300 312  
 FT DISULFID 316 332  
 FT DISULFID 335 339  
 FT DISULFID 339 339  
 FT DISULFID 513 522  
 FT DISULFID 517 530  
 FT DISULFID 533 542  
 FT DISULFID 546 562  
 FT DISULFID 565 578  
 FT DISULFID 569 586  
 FT DISULFID 589 598  
 FT DISULFID 602 625  
 FT DISULFID 628 636  
 FT DISULFID 632 644  
 FT MOD\_RES 632 644  
 FT MOD\_RES 1141 1141  
 FT MOD\_RES 1250 1250  
 FT CARBOHYD 68 68  
 FT CARBOHYD 188 188  
 FT CARBOHYD 260 260  
 FT CARBOHYD 532 532  
 FT CARBOHYD 573 573  
 FT CARBOHYD 631 631

FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).  
 SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;  
 Query Match 80.7%; Score 954; DB 1; Length 1257;  
 Best Local Similarity 81.6%; Pred. No. 1e-50; Indels 0; Gaps 0;  
 Matches 177; Conservative 8; Mismatches 32;  
 QY 1 GAGGMVHRRSSSTRSGGDLTLGLEPSEEAARSPPLASEGASDPFDGDMGAANK 60  
 DB 1040 GTSTAHRRRRSSSTRSGGDLTLGLEPSEEAARSPPLASEGASDPFDGDMGAANK 1099  
 QY 61 LQSLPHDPSPDLQRYSDPLVPPLPSETDGYVAPLTCSPOPEYVQNPVRRQPPSPREGPL 120  
 DB 1100 LQSLSPHDPSPDLQRYSDPLVPPLPSETDGYVAPLTCSPOPEYVQNPVRRQPPSPREGPL 1159  
 QY 121 PAARPAATLERAKTSPGNGVYKDVFAFGAVENNEYTLPOGGAPOPHPPAPSPAF 180  
 DB 1160 PVPAPGATLERKRTLSPGNGVYKDVFAFGAVENNEYTLVPRAGTASPPHSPAPSPAF 1219  
 QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
 DB 1220 DNLVYWDQDQSEEGPPSPNEGTPTAENPEYLGIDVP 1256  
 RESULT 4  
 ID ERB2\_MOUSE STANDARD; PRT; 245 AA.  
 AC P70424; Q61525;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 (BC 2.7.1.112) (p185erbB2)  
 DE (NEU proto-oncogene) (C-erbB-2) (Fragments).  
 GN ERB2 OR NEU.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RC STRAIN=CD-1; TISSUE=Uterus;  
 RX MEDLINE=97200814; PubMed=9048643;  
 RA Kim J., Dey S.K., Das S.K.;  
 RT "Differential expression of the erbB2 gene in the perimplantation  
 RT mouse uterus: potential mediator of signaling by epidermal growth  
 RL factor-like growth factors.";  
 RN Endocrinology 138:1328-1337(1997).  
 RP SEQUENCE OF 150-245 FROM N.A.  
 RX MEDLINE=96069911; PubMed=7569796;  
 RA Moscoso L.M., Chu G.C., Gautam M., Noakes P.G., Merlie J.P.,  
 RA Sanes J.R.;  
 RT "Synapse-associated expression of an acetylcholine receptor-inducing  
 RT protein, ARIA/mergulin, and its putative receptors, ErbB2 and ErbB3,  
 RT in developing mammalian muscle.";  
 RL Dev. Biol. 172:158-169(1995).  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,  
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
 CC ALPHA AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS  
 CC (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN UTERINE EPITHELIAL  
 CC CELLS. IN THE MUSCLE, EXPRESSION LOCALIZES TO THE SYNAPTIC SITES  
 CC OF MUSCLE FIBERS.  
 CC -1- DEVELOPMENTAL STAGE: ON DAYS 1-4 OF PREGNANCY, ERBB2 IS DETECTED  
 CC PRIMARILY IN EPITHELIAL CELLS, THE DAY 1 UTERUS SHOWING THE  
 CC HIGHEST ACCUMULATION. ON DAY 5, THE EPITHELIUM AND THE  
 CC DECIDUALIZING STROMAL CELLS AROUND THE IMPLANTING BLASTOCYST  
 CC EXHIBIT ACCUMULATION OF THIS RECEPTOR. ON DAYS 6-8, THE EXPRESSION

CC PERSISTS IN THE EPITHELIUM AT BOTH THE IMPLANTATION AND  
 CC INTERIMPLANTATION SITES IN ADDITION TO MODEST LEVELS IN THE  
 CC SECONDARY DECIDUAL ZONE. ON DAYS 7 AND 8, ACCUMULATION IS ALSO  
 CC PROMINENT IN THE TROPHOBLASTIC GIANT CELLS.  
 CC -1- P7M: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U71126; AAB17380.1; -  
 CC EMBL: L47239; AAG93532.1; -  
 CC HSSP: P11362; 1FGK.  
 CC MGD: MGI:95410; ErbB2.  
 CC InterPro: IPR000719; ErbB2.  
 CC InterPro: IPR004040; STY\_Pkinase.  
 CC InterPro: IPR001245; TYR\_Pkinase.  
 CC PRINTS: PR00109; TYRKINASE.  
 CC PRODOM: PD000001; Euk\_Pkinase: 1.  
 CC SMART: SM00221; STYKC: 1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP: PARTIAL.  
 CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR: 1.  
 CC PROSITE: PS50011; PROTEIN\_KINASE\_DOM: 1.  
 CC Transmembrane; Glycoprotein; Multigene family; Receptor; Transferase;  
 CC Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
 CC  
 CC FT DOMAIN 1 1  
 CC FT ACT\_SITE <1 149 PROTEIN KINASE.  
 CC FT NON\_CONS 61 61 BY SIMILARITY.  
 CC FT NON\_TER 149 150  
 CC FT SEQUENCE 245 245  
 CC  
 CC SQ SEQUENCE 245 AA; 26927 MW; 0F763F0363DEF1C CRC64;  
 CC  
 CC Query Match 23.9%; Score 283; DB 1; Length 245;  
 CC Best Local Similarity 79.4%; Pred. No. 6,2e-11;  
 CC Matches 54; Conservative 2; Mismatches 12; Indels 0; Gaps 0;  
 CC  
 CC QY 1 GAGGHHRRSSSTRSGGDLTGLEPSEEPSPPLAPSEAGSDVFDGLGKAAG 60  
 CC Db 177 GTGSTAHRHRSSSRSSARSGGELTGLEPSEEPSPPLAPSEAGSDVFDGLAVGYTKG 236  
 CC  
 CC QY 61 LQSLPTHD 68  
 CC Db 237 LQSLSPHD 244  
 CC  
 CC RESULT 5  
 CC EGF\_R\_HUMAN STANDARD; PRT; 1210 AA.  
 CC AC P00533; P06268; Q14225; Q9UMD7; Q9UMD8; Q9UMG5; Q92795; O00732;  
 CC AC O00683; Q9BZS2; Q9BZC9; Q9HZC9; Q9HZX1; Q9H3C9;  
 CC DT 21-JUL-1986 (Ref. 01, Created)  
 CC DT 01-NOV-1997 (Ref. 35, Last sequence update)  
 CC DT 15-JUN-2002 (Ref. 41, Last annotation update)  
 CC DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor  
 CC protein-tyrosine kinase ErbB-1).  
 CC GN EGF\_R OR ERBB1.  
 CC OS Homo sapiens (Human).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 CC OX NCL\_TaxID=9606;  
 CC RN 11  
 CC RP SEQUENCE FROM N.A. (ISOFORM 1).  
 CC RX MEDLINE=84219729; PubMed=6328312;  
 CC RX Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
 CC Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,  
 CC Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;  
 CC RA Human epidermal growth factor receptor cDNA sequence and aberrant

RT expression of the amplified gene in A431 epidermoid carcinoma cells.";  
 RL Nature 309:418-425(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Placenta;  
 RX MEDLINE=95382957; PubMed=7654368;  
 RA Ilekis J.V., Stark B.C., Scoccia B.;  
 RT "Possible role of variant RNA transcripts in the regulation of  
 RT epidermal growth factor receptor expression in human placenta.";  
 RL Mol. Reprod. Dev. 41:149-156(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Placenta;  
 RX MEDLINE=97078686; PubMed=9918811;  
 RA Reiter J.L., Mahle N.J.;  
 RT "A 1.8 kb alternative transcript from the human epidermal growth  
 RT factor receptor gene encodes a truncated form of the receptor.";  
 RL Nucleic Acids Res. 24:4050-4056(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Placenta;  
 RX MEDLINE=97256547; PubMed=9103388;  
 RA Ilekis J.V., Garfili J., Niederberger C., Scoccia B.;  
 RT "Expression of a truncated epidermal growth factor receptor-like  
 RT protein (TEGR) in ovarian cancer.";  
 RL Gynecol. Oncol. 65:36-41(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
 RC TISSUE=Placenta;  
 RX MEDLINE=21100872; PubMed=11161793;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Scheel Sinclair C., Peersall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,  
 RA Mahle N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative EGF receptor transcripts encoding truncated receptor  
 RT isoforms.";  
 RL Genomics 71:1-20(2001).  
 RN [6]  
 RP SEQUENCE OF 575-687 FROM N.A.  
 RA Reiter J.L., Threadgill D.W., Danielsen A.J., Scheel C.M.,  
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
 RA Mahle N.J.;  
 RT "Human and mouse alternative EGF receptor transcripts encoding only the  
 RT extracellular domain of the receptor.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 713-924 FROM N.A.  
 RX MEDLINE=84196372; PubMed=6326261;  
 RA Lin C.R., Chen W.S., Kruller W., Stolarsky L.S., Weber W.,  
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;  
 RT "Expression cloning of human EGF receptor complementary DNA: gene  
 RT amplification and three related messenger RNA products in A431  
 RT cells.";  
 RL Science 224:843-848(1984).  
 RN [8]  
 RP SEQUENCE OF 150-962 FROM N.A.  
 RX MEDLINE=84245835; PubMed=6330563;  
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
 RA Roe B.A., Merlino G.T., Pastan I.;  
 RT "Human epidermal growth factor receptor cDNA is homologous to a  
 RT variety of RNAs overproduced in A431 carcinoma cells.";  
 RL Nature 309:806-810(1984).  
 RN [9]  
 RP SEQUENCE OF 1028-1210 FROM N.A.  
 RX MEDLINE=85046483; PubMed=6093780;  
 RA Slamen F.A., Gope M.L., Schultz T.Z., Wright D.A., Carpenter G.,  
 RA O'Malley B.W.;  
 RT "Isolation of an evolutionarily conserved epidermal growth factor  
 RT receptor cDNA from human A431 carcinoma cells.";  
 RL Biochem. Biophys. Res. Commun. 124:125-132(1984).  
 RN [10]  
 RP SEQUENCE OF 1-29 FROM N.A.

RA MEDLINE-88217333; PubMed-3329716;  
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
 RA Waterfield M.D.;  
 RT "The human EGF receptor gene: structure of the 110 kb locus and  
 RT identification of sequences regulating its transcription.";  
 RL Oncogene Res. 1:375-396(1987).  
 RN [111]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE-91107677; PubMed-1988448;  
 RA Haley J.D., Waterfield M.D.;  
 RT "Contributory effects of de novo transcription and premature  
 RT transcript termination in the regulation of human epidermal growth  
 RT factor receptor proto-oncogene RNA synthesis.";  
 RL J. Biol. Chem. 266:1746-1753(1991).  
 RN [112]  
 RP SEQUENCE OF 1-29 FROM N.A.  
 RX MEDLINE-85270438; PubMed-2991899;  
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;  
 RT "Characterization and sequence of the promoter region of the human  
 RT epidermal growth factor receptor gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
 RN [113]  
 RP SEQUENCE OF 540.  
 RA Kohda D.;  
 RL Submitted (SEP-1997) to the SWISS-PROT data bank.  
 RN [114]  
 RP RECEPTOR ACTIVITY.  
 RX MEDLINE-84191554; PubMed-6325948;  
 RA Mroczkowski B., Mosig G., Cohen S.;  
 RT "ATP-stimulated interaction between epidermal growth factor receptor  
 RT and supercoiled DNA".  
 RL Nature 309:270-273(1984).  
 RN [115]  
 RP PHOSPHORYLATION.  
 RX MEDLINE-89278137; PubMed-2543678;  
 RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,  
 RA Howk R., Glynn D., Ullrich A., Schlessinger J.;  
 RT "All autophosphorylation sites of epidermal growth factor (EGF)  
 RT receptor and HER2/neu are located in their carboxyl-terminal tails.  
 RT Identification of a novel site in EGF receptor".  
 RL J. Biol. Chem. 264:10667-10671(1989).  
 RN [116]  
 RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
 RP ASN-528.  
 RX MEDLINE-96398132; PubMed-8962717;  
 RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;  
 RT "Analysis of the glycosylation patterns of the extracellular domain of  
 RT the epidermal growth factor receptor expressed in Chinese hamster  
 RT ovary fibroblasts".  
 RL Glycoph Factors 13:121-132(1996).  
 RN [117]  
 RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND  
 RP ASN-603.  
 RX MEDLINE-20198209; PubMed-10731668;  
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;  
 RT "Characterization of the N-oligosaccharides attached to the atypical  
 RT Asn-X-Cys sequence of recombinant human epidermal growth factor  
 RT receptor".  
 RL J. Biochem. 127:65-72(2000).  
 RN [118]  
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
 RX MEDLINE-98225196; PubMed-9556602;  
 RA Abe Y., Otake M., Inagaki F., Lax I., Schlessinger J., Kohda D.;  
 RT "Disulfide bond structure of human epidermal growth factor receptor".  
 RL J. Biol. Chem. 273:11150-11157(1998).  
 RN [119]  
 RP REVIEW.  
 RX MEDLINE-87297456; PubMed-3039909;  
 RA Carpenter G.;  
 RT "Receptors for epidermal growth factor and other polypeptide  
 RT mitogens".  
 RL Annu. Rev. Biochem. 56:881-914(1987).  
 CC -I- FUNCTION: Receptor for EGF, but also for other members of the EGF

CC family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding  
 CC EGF-like growth factor, gp30 and vaccine virus growth factor. Is  
 CC involved in the control of cell growth and differentiation.  
 CC -I- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.  
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
 CC tyrosine phosphate.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is  
 CC secreted.  
 CC -I- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/  
 CC truncated isoform/TEGFR, 3/p110 and 4; are produced by  
 CC alternative splicing.  
 CC -I- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also  
 CC expressed in ovarian cancers.  
 CC -I- MISCELLANEOUS: Binding of EGF to the receptor leads to  
 CC dimerization, internalization of the EGF-receptor complex,  
 CC induction of the tyrosine kinase activity, stimulation of cell DNA  
 CC synthesis, and cell proliferation.  
 CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC DR EMBL; X00588; CA025240.1; -  
 CC DR EMBL; U95089; AAB53063.1; -  
 CC DR EMBL; U48723; AAC50802.1; -  
 CC DR EMBL; U48723; AAC50804.1; -  
 CC DR EMBL; U48724; AAC50796.1; -  
 CC DR EMBL; U48725; AAC50797.1; -  
 CC DR EMBL; U48726; AAC50798.1; -  
 CC  
 CC Query Match 13.9%; Score 164; DB 1; Length 1210;  
 CC Best Local Similarity 30.5%; Pred. No. 0.004%; Gaps 11;  
 CC Matches 62; Conservative 24; Mismatches 69; Indels 48; Gaps 11;  
 CC  
 CC 29 SEEEAPRSLPAPSGAGSDVFDGDLGMAKGLQSLPTHTDPSPLQRYSEDPVLPSET- 87  
 CC 1025 SSPSTSTPLSLSLASATSN--NSTVACIDRNGQSCIRKEDSFQRRSSPTGLTSDSI 1082  
 CC  
 CC 88 -DGVADPLTCSPOPEYVQNDVDRPQPPSPREGLPAPRPGATLIERAKTISPCKNGVYKD 146  
 CC DB 1083 DDFVL-----PVPEYVNIQ--SVPRPAGSVQNPVYHNPQPLNP-----APSRDPHYQD 1127  
 CC QY 147 VFAPFGAVENPEYL-TPGGAAQPPHPAPFAFDLXYWDQ-----DP----- 190  
 CC DB 1128 --PHSTAVGNPEYLTNYQ-----PTCVNSTFDSPAHMAQSGSHQISLDNDPYQODF 1176  
 CC  
 CC QY 191 -PERGAPSTFKGTPTANPEYL 212  
 CC DB 1177 FPKFAKNGIFKKS-TAENAEYL 1198  
 CC  
 CC RESULT 6  
 CC ERB4\_HUMAN  
 CC ID ERB4\_HUMAN STANDARD; PRT; 1308 AA.  
 CC AC 015303;  
 CC DT 15-DEC-1998 (Rel. 37, Created)  
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)  
 CC DE (p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).  
 CC GN ERB4 OR HER4  
 CC OS Homo sapiens (human).  
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC OX NCBI\_TaxID=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A. (ISOFORM JM-A).  
 CC RC TISSUE-Breast carcinoma;



RX MEDLINE-93189574; PubMed-8383326;  
 RA Plozman G.D., Culouscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,  
 RA Foy L., Neuhauer M.G., Shoyab M.;  
 RT "Ligand-specific activation of HER4/p180erbB4, a fourth member of the  
 RT epidermal growth factor receptor family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).  
 RC TISSUE-Fetal brain:  
 RA MEDLINE-97476287; PubMed-9334263;  
 RA Elenius K., Cortas G., Paul S., Choi C.J., Rio C., Plozman G.D.,  
 RA Klagsbrun M.;  
 RT "A novel junctional domain isoform of HER4/erbB4. Isoform-specific  
 RT tissue distribution and differential processing in response to  
 RT phorbol ester".  
 RL J. Biol. Chem. 272:26761-26768(1997).  
 CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-  
 CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND  
 CC NTRK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.  
 CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB  
 CC RECEPTORS (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: JM-A (SHOWN HERE) AND JM-B;  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER  
 CC FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED  
 CC BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND  
 CC NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN  
 CC CEREELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,  
 CC KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREELLUM,  
 CC PITUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,  
 CC LUNG, SALIVARY GLAND, AND PANCREAS.  
 CC -1- PFM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L07868; AAB59446.1; -;  
 DR HSSP: P11362; IFGK.  
 DR Genew: HGNC:3432; ERBB4.  
 DR MIM: 600543; -;  
 DR InterPro: IPR000494; EGFR\_Ldomain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR InterPro: IPR004019; YLP\_moclit.  
 DR Pfam: PF00069; pkinase.1.  
 DR Pfam: PF00757; Furin-like.1.  
 DR Pfam: PF01030; Recep\_Ldomain.2.  
 DR Pfam: PF02757; YLP.2.  
 DR ProDom: PD000001; Euk\_Pkinase.1.  
 DR SMART: SM00261; FU; 4.  
 DR SMART: SM00219; TyrcK.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Alternative splicing.  
 FT SIGNAL 1 25 POTENTIAL.  
 FT CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4.  
 FT DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 652 675 POTENTIAL.

FT DOMAIN 676 1308 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 186 334 CYS-RICH.  
 FT DOMAIN 496 633 CYS-RICH.  
 FT DOMAIN 718 985 PROTEIN KINASE.  
 FT NP\_BIND 721 732 ATP (BY SIMILARITY).  
 FT BINDING 751 751 ATP (BY SIMILARITY).  
 FT ACT\_SITE 843 843 BY SIMILARITY.  
 FT DISULFID 189 197 BY SIMILARITY.  
 FT DISULFID 249 258 BY SIMILARITY.  
 FT DISULFID 262 289 BY SIMILARITY.  
 FT DISULFID 293 304 BY SIMILARITY.  
 FT DISULFID 308 323 BY SIMILARITY.  
 FT DISULFID 326 330 BY SIMILARITY.  
 FT DISULFID 503 512 BY SIMILARITY.  
 FT DISULFID 507 520 BY SIMILARITY.  
 FT DISULFID 523 532 BY SIMILARITY.  
 FT DISULFID 536 552 BY SIMILARITY.  
 FT DISULFID 555 569 BY SIMILARITY.  
 FT DISULFID 559 577 BY SIMILARITY.  
 FT DISULFID 580 589 BY SIMILARITY.  
 FT DISULFID 593 614 BY SIMILARITY.  
 FT DISULFID 617 625 BY SIMILARITY.  
 FT DISULFID 621 633 BY SIMILARITY.  
 FT MOD\_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 495 495 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 548 548 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 626 648 N-GLYCOSYLATION (AUTO-) (BY SIMILARITY).  
 FT VARSPLIC 648 648 NGPTSHDCIYPMVGHSTLPDHA -> IGSSIEDICIGLMD  
 FT (IN ISOFORM JM-B).  
 FT SO SEQUENCE 1308 AA; 146807 MW; 5E4AE80985D88761 CRC64;  
 Query Match 13.38; Score 157; DB 1; Length 1308;  
 Best Local Similarity 24.28; Pred. No. 0.012;  
 Matches 62; Conservative 23; Mismatches 75; Indels 96; Gaps 10;  
 QY 16 RSGGDLTLGLE-PSEEAAPRSLAP-SEGASGVDPDGLGCAKAGLSLPTHPSPLO 73  
 DB 1067 RDGFAAEGVSVYPRAPRTSTLPEAPVAGATAETFDSCCGNTRAKKVAAPVQEDSSSTQ 1126  
 QY 74 RYSDPTVPLPS-----ETDGYAAPLTCSPQPEYVQNPQVRRQPPSPRSGPLPAABA 126  
 DB 1127 RYSADPTVFAPERSPRGELDEGVTWPRDKRQGYLNPVE-----ENPFVSR- 1175  
 QY 127 GATLERAKTLSPGKNGVYKDFAFGCAVENEPEYLTPOGCAAPQPIPPA- 175  
 DB 1176 -----KNGDLQ-----ALDNEPYHNASNG-----PPRAEDRYNEPEYL 1209  
 QY 176 -----FSPADNLYYMPDQDPERGA--PSPNFGTPT----- 205  
 DB 1210 NTPANTLCKAEYLNKNNIISMPEKAKKAPDNDPYNHSLPPRSTLOHPDYLCTSTKYRYK 1269  
 QY 206 -----AENPEYL 212  
 DB 1270 QNGRTRPIVAENPEYL 1285  
 RESULT 7

ERB4\_RAT STANDARD; PRT: 1308 AA.  
 ID ERB4\_RAT  
 AC Q62956; Q922N7;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).  
 GN ERB4 OR TYRO-2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID:10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=98221155; PubMed=9553078;  
 RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,  
 RA Marchionni M.A., Kelly R.A.;  
 RT "Neuregulins promote survival and growth of cardiac myocytes.  
 RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult  
 RT ventricular myocytes.";  
 RL J. Biol. Chem. 273:10261-10269(1998).  
 RN [2]  
 RP SEQUENCE OF 848-901 FROM N.A.  
 RC TISSUE=Sciatic nerve;  
 RX MEDLINE=91223560; PubMed=20253425;  
 RA Lai C., Lemke G.;  
 RT "An extended family of protein-tyrosine kinase genes differentially  
 RT expressed in the vertebrate nervous system.";  
 RL Neuron 6:691-704(1991).  
 RN [3]  
 RP SEQUENCE OF 1031-1198 FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Spinal cord;  
 RX MEDLINE=97184212; PubMed=9030624;  
 RA Carroll S.L., Miller M.L., Frohnert P.W., Kim S.S., Corbett J.A.;  
 RT "Expression of neuregulins and their putative receptors, ErbB2 and  
 RT ErbB3, is induced during Wallerian degeneration.";  
 RL J. Neurosci. 17:1642-1659(1997).  
 CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-  
 CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND  
 CC NTAR. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.  
 CC NOT ACTIVATED BY EGF, TGF- $\alpha$ , AND AMPHIREGULIN (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERB  
 CC RECEPTORS (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING  
 CC NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS  
 CC OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE  
 CC RETICULAR NUCLEUS OF THE THALAMUS, VERY LOW LEVELS IN KIDNEY, AND  
 CC HEART.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF041838; AAD08899.1; -  
 CC EMBL: U52531; AAC53051.1; -  
 DR HSPB, P11362; 1FCR.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00757; Furin-like; 1.

DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR Pfam: PF02757; YLP; 2.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; F0; 4.  
 DR SMART: SM00219; TYRKC; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_DGM; 1.  
 KW Transmembrane, Glycoprotein, Multigene family, Receptor, Signal;  
 KW Transferrase, Tyrosine-protein kinase, ATP-binding, Phosphorylation.  
 FT SIGNAL 1 25  
 FT CHAIN 1 25  
 FT DOMAIN 26 1308  
 FT TRANSMEM 652 675  
 FT DOMAIN 676 1308  
 FT DOMAIN 186 334  
 FT DOMAIN 496 633  
 FT DOMAIN 718 985  
 FT NP\_BIND 724 732  
 FT BINDING 751 751  
 FT ACT\_SITE 843 843  
 FT DISULFID 189 197  
 FT DISULFID 193 205  
 FT DISULFID 213 221  
 FT DISULFID 217 229  
 FT DISULFID 230 238  
 FT DISULFID 234 246  
 FT DISULFID 249 258  
 FT DISULFID 262 289  
 FT DISULFID 293 304  
 FT DISULFID 308 323  
 FT DISULFID 326 330  
 FT DISULFID 503 512  
 FT DISULFID 507 520  
 FT DISULFID 523 532  
 FT DISULFID 536 552  
 FT DISULFID 555 569  
 FT DISULFID 559 577  
 FT DISULFID 580 589  
 FT DISULFID 593 614  
 FT DISULFID 617 625  
 FT DISULFID 621 633  
 FT MOD\_RES 1162 1162  
 FT MOD\_RES 1188 1188  
 FT MOD\_RES 1258 1258  
 FT MOD\_RES 1284 1284  
 FT CARBOHYD 138 138  
 FT CARBOHYD 174 174  
 FT CARBOHYD 253 253  
 FT CARBOHYD 358 358  
 FT CARBOHYD 410 410  
 FT CARBOHYD 473 473  
 FT CARBOHYD 495 495  
 FT CARBOHYD 348 348  
 FT CARBOHYD 576 576  
 FT CARBOHYD 620 620  
 FT CONFLICT 1062 1062  
 FT CONFLICT 1080 1082  
 SQ SEQUENCE 1308 AA; 146957 MW; D944B0996A08B41 CRC64;  
 Query Match 12.9%; Score 152; DB 1; Length 1308;  
 Best Local Similarity 24.2%; Pred. No. 0.024;  
 Matches 57; Conservative 23; Mismatches 60; Indels 96; Gaps 9;  
 QY 34 PRSPLAPSEAGSDVFDGDLGKGAAGLGLPHRDSPLDRYSDDPRVLP-----E 86  
 DB 1089 PEAPVA--QGATFEMPDSCNCTLRKPVVPHVQESSTQRYSDAPVFAERNPRAELD 1146  
 QY 87 TDGYVAPLTCSPQPEVYNQPDVVRPPSPREGGLPARPAGATLERAKTLSPGKNGVKD 146  
 DB 1147 EBGYTMPMDHKPKQRYLNPVE-----ENPFVSHR-----KNGDLO 1181

```

OY 147 VFAAGAVENPEYLTPOGGAAPQPPPPA----- 175
DB 1182 -----ALDNEFYHSASSG-----PKRADEYNEPLYLNTFNALGNAEYKNSLSLV 1229
OY 176 ---FSPAFLNLYWDDPPPERGA---PPSTFKGPT-----AENPEYL 212
DB 1230 PEKAKKAFNDNDYNNHSLPRSTLQHPDYLQDYSTKYFYKONGHRIPIVANPEYL 1285

RESULT 8
SFPO_HUMAN
ID SFPO_HUMAN STANDARD; PRT; 707 AA.
AC P23246; P30808;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Splicing factor, proline- and glutamine-rich (polypyrimidine tract-
DE binding protein-associated splicing factor) (PTB-associated splicing
DE factor) (PSF) (DNA-binding p52/p100 complex, 100 kDa subunit).
OS SFPO OR PSF.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.
RC TISSUE=Fetal brain;
RA MEDLINE=93194059; PubMed=8449401;
RA Patton J.G., Porro E.B., Galceran J., Tempst P., Nadal-Ginard B.;
RT "Cloning and characterization of PSF, a novel pre-mRNA splicing
RT factor."
RL Genes Dev. 7:393-406(1993).
RN [2]
RP SEQUENCE OF 312-707 FROM N.A.
RC TISSUE=Fetal skeletal muscle;
RX MEDLINE=90091812; PubMed=2480877;
RA Gower H.J., Moore S.E., Dickson G., Elsom V.L., Nayak R., Walsh F.S.;
RT "Cloning and characterization of a myoblast cell surface antigen
RT defined by 24 Id5 monoclonal antibody."
RL Development 105:723-731(1997).
RN [3]
RP SEQUENCE OF 48-68 AND 213-246.
RX MEDLINE=93176127; PubMed=8439294;
RA Zhang W.-W., Zhang L.-X., Busch R.K., Fairies J., Busch H.;
RT "Purification and characterization of a DNA-binding heterodimer of 52
RT and 100 kDa from HeLa cells."
RL Biochem. J. 290:267-272(1993).
CC -1- FUNCTION: ESSENTIAL PRE-MRNA SPLICING FACTOR REQUIRED EARLY IN
CC SPLICOSOME FORMATION. BINDS TO THE MAMMALIAN POLYPRIMIDINE
CC TRACTS. FORMS A COMPLEX WITH THE POLYPRIMIDINE TRACT-BINDING
CC PROTEIN (PTB). SEEMS TO ALSO BIND DNA.
CC -1- SUBUNIT: HETEROTETRAMER OF TWO 52 KDA AND TWO 100 KDA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; A LONG FORM (SHOWN
CC HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM)
CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE MYOBLAST CELL
CC SURFACE ANTIGEN 24.ID5 AND A POSSIBLE MEMBRANE-BOUND PROTEIN
CC ECTOIMINSE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70944; CA50283.1; -
DR EMBL: X16850; CA34747.1; -
DR PIR: A43557; A43557.
DR PIR: S29770; S29770.

```

```

DR HSSP; P11940; ICVJ.
DR SWISS-2DPAGE; P23246; HUMAN.
DR Genew: HGNC:10774; SFPO.
DR MIM; 605199; -.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Nuclear protein; RNA-binding; DNA-binding; mRNA splicing; Repeat;
KW Alternative splicing.
FT DOMAIN 297 369 RNA-BINDING (RRM) 1.
FT DOMAIN 371 452 RNA-BINDING (RRM) 2.
FT DOMAIN 9 27 3 X 3 AA REPEATS OF R-G-G.
FT REPEAT 19 21 1.
FT REPEAT 25 27 2.
FT REPEAT 10 15 3.
FT DOMAIN 10 15 GLN/GLU/PRO-RICH.
FT DOMAIN 20 27 POLY-GLY.
FT DOMAIN 56 65 POLY-PRO.
FT DOMAIN 67 71 POLY-GLN.
FT DOMAIN 95 98 POLY-GLN.
FT DOMAIN 99 103 POLY-PRO.
FT DOMAIN 184 188 POLY-PRO.
FT DOMAIN 571 574 POLY-ARG.
FT DOMAIN 613 616 POLY-GLY.
FT DOMAIN 635 641 POLY-GLY.
FT VARSPLIC 663 707 POLY-GLY.
FT CONFLICT 243 243 RTERFGGAGPVGCGGCPGRCGCTPAGTCGRGREGC
FT FT PNKRPRF -> VRMIDVG (IN SHORT ISOFORM).
FT FT G -> R (IN REF. 3).
SQ SEQUENCE 707 AA; 76149 MW; 6D8D5EA35E235847 CRC64;

Query Match 12.3%; Score 145.5; DB 1; Length 707;
Best Local Similarity 26.6%; Pred. No. 0.033;
Matches 58; Conservative 13; Mismatches 82; Indels 65; Gaps 10;

OY 1 GAGGAVHRRHSSSTRSGGDLTLGLEPSEEEAPRSLAPSEGAQSDVFDGLGGAAG 60
DB 11 GGGGGRHRR-----GGGGRGLHFR-----SPPGAGLNNRPGMKRPGQSS 54
OY 61 LOSLETHDPSPLQRYSE-DPTVPLPSETDGYVAPLTCSPOPEYVNPDPVRQPPSPRE- 117
DB 55 GPKPRPIPPPHQOQQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 107
OY 118 -----GPIPARPACATIERAKTILSPKNGVYKDVAFAGAVENPEYLTPOGGAAPQ 169
DB 108 KPYVAQGPAPGVGSAPASSAPAPPTPSGA-----PPGSGPGPT 150
OY 170 PHPPAFSPAFDNLVYWDPPPERGAPST--FKGPT 205
DB 151 PTPPAVTSA-----PP--GAPPTPPSSGVPT 176

RESULT 9
ERBB_ALV
ID ERBB_ALV STANDARD; PRT; 634 AA.
AC P00534;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transforming protein erbB (EC 2.7.1.112).
GN V-ERBB.
OS Avian leukosis virus.
OC Viruses; Retroviridae; Alpharetrovirus.
OX NCBI_TaxID=11864;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85228222; PubMed=2988784;
RA Nilsen T.W., Maroney P.A., Goodwin R.G., Rottman F.M.,
RA Crittenden L.B., Raines M.A., Kung H.-U.;
RT "c-erbB activation in ALV-induced erythroblastosis: novel RNA
RT processing and promoter insertion result in expression of an

```

RT amino-truncated EGF receptor."  
 RL Cell 41:719-726(1985).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS PRODUCED BY ERYTHROLEUKEMIA CELLS  
 CC IN CHICKS WHICH HAVE BEEN INJECTED WITH THE AVIAN LEUKOSIS VIRUS  
 CC AND SUBSEQUENTLY DEVELOP LEUKEMIA. THE LEUKEMIAS ARE INDUCED BY  
 CC THE INSERTION OF RETROVIRAL PROMOTER ELEMENTS IN OR NEAR THE  
 CC C-ERBB GENE, WHICH LEADS TO C-ERBB ACTIVATION.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ENV-ERBB  
 CC PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M10066; AAA48763.1; ALT\_INIT.  
 CC PIR: A00643; TVCHLV.  
 CC PIR: B00643; TVFVLV.  
 CC HSP: P1352; 1FCR.  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR001245; Tyr\_pkinase.  
 CC Pfam: PF00069; pkinase; 1.  
 CC PRINTS: PR00109; TYRKINASE.  
 CC PRODOM: PD000001; Euk.pkinase; 1.  
 CC SMART: SM00219; TyrcK; 1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC KW Transferrase: Tyrosine-protein kinase; ATP-binding; Oncogene;  
 CC Glycoprotein; Phosphorylation.  
 CC FT DOMAIN 132 399  
 CC NP\_BIND 138 146 ATP (BY SIMILARITY).  
 CC BINDING 165 165 ATP (BY SIMILARITY).  
 CC ACT\_SITE 257 257 BY SIMILARITY.  
 CC SEQUENCE 634 AA; 70891 MW; E705E33A0BE01FCC CRC64;  
 SQ  
 Query Match 11.7%; Score 138; DB 1; Length 634;  
 Best Local Similarity 24.5%; Pred. No. 0.085;  
 Matches 54; Conservative 25; Mismatches 59; Indels 82; Gaps 10;  
 QY 35 RSFL-----APSEGAGSDVFDGDLGMAKGLQSLPTHPDSPLOQRYSEDPVLPSEFT-- 87  
 DB 451 RFLSSLSLSTNSNSATNCID-----RNGGCHYVREDSFYQRTSSDPTGFLRESID 502  
 QY 88 DGVAVPLTCSPOPEYVNPDPVRPQPSRPGPLPAARPGATLERAKTISPGKNGVAKDV 147  
 DB 503 DGFLL-----PAPEYVNO--LMPKPS-----TAMVNOQIYNNI 533  
 QY 148 F-----AFGAVENPEYLTPOGGAAPQHPHFAFADLYWDQ----- 188  
 DB 534 SLTAIKLPMDSRYQNSHSTAVDNPETL-----NTNOSPLAKIVFESSPYWIOSGNH 585  
 QY 189 ----DPPE-----RGAPPSTFKGTPTAENPEYLGLDVP 217  
 DB 586 QIULDNPDYQODFLPMETKPNGLKVPAAENPEYLVAAAP 625  
 RESULT 10  
 PREL\_HUMAN  
 ID PREL\_HUMAN STANDARD: PRT: 331 AA.  
 AC P04280:  
 DT 20-MAR-1987 (rel. 04, Created)  
 DT 13-AUG-1987 (rel. 05, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Salivary proline-rich protein precursor (Clones CP3, CP4 and CP5)  
 DE [Contains: Basic peptide IB-6; Peptide P-H].  
 GN PRB1.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 OX  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85289325; PubMed=2993301;  
 RA Maeda N., Kim H.-S., Azen E.A., Smithies O.;  
 RT "Differential RNA splicing and post-translational cleavages in the  
 RT human salivary proline-rich protein gene system.";  
 RL J. Biol. Chem. 260:11123-11130(1985).  
 RN  
 RP SEQUENCE OF 214-331.  
 RX MEDLINE=86243355; PubMed=3521730;  
 RA Kaufman D., Hofmann T., Bennick A., Keller P.;  
 RT "Basic proline-rich proteins from human parotid saliva: complete  
 RT covalent structures of proteins IB-1 and IB-6.";  
 RL Biochemistry 25:2387-2392(1986).  
 RN  
 RP SEQUENCE OF 276-331.  
 RX MEDLINE=84161824; PubMed=6671974;  
 RA Saltoh E., Isemura S., Sanada K.;  
 RT "Further fractionation of basic proline-rich peptides from human  
 RT parotid saliva and complete amino acid sequence of basic proline-rich  
 RT peptide P-H.";  
 RL J. Biochem. 94:1991-1997(1983).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: K03204; AAA60185.1; -  
 CC DR EMBL: K03205; AAA60186.1; -  
 CC DR EMBL: K03206; AAA60187.1; -  
 CC DR PIR: A03291; PIH0B6.  
 CC DR PIR: C25372; C25372.  
 CC DR Genew: HGNC:9337; PRB1.  
 CC KW MIM: 180989; -  
 CC REM: Parotid gland; Multigene family; Saliva; Signal.  
 CC FT SIGNAL 1 16  
 CC FT CHAIN 214 331 PEPTIDE IB-6.  
 CC FT CHAIN 276 331 PEPTIDE P-H.  
 CC FT VARIANT 106 238 MISSING (IN CLONE CP-4).  
 CC FT VARIANT 106 258 /FTID-VAR.005561.  
 CC FT VARIANT 106 258 MISSING (IN CLONE CP-5).  
 CC FT CONFLICT 276 276 /FTID-VAR.005562.  
 CC FT CONFLICT 276 276 A -> S (IN REF. 2 AND 3).  
 CC SQ SEQUENCE 331 AA; 32596 MW; 3F481FF8EBA39751 CRC64;  
 Query Match 11.5%; Score 135.5; DB 1; Length 331;  
 Best Local Similarity 25.5%; Pred. No. 0.064;  
 Matches 64; Conservative 17; Mismatches 76; Indels 94; Gaps 13;  
 QY 11 RSSSTRSGGDLTLGLEPSEEARSPPLAP---SEG---AGSDVFDGDLGMAKGLQSL 64  
 DB 86 KSRSPRSPEPK-PQGRPPPGGNGQPGPPPPKPGQGRPPGKPKPGPPPPKPGQ---- 140  
 QY 65 PTHDSPLOQRYSEDPTVPPLPSETDGYVAPLTCSPOPEYVNPDPVRPQPS-----PR--- 116  
 DB 141 ----PPPGDKSQSPRSP-PGKPGQ-----PPPGGNGQPGPPPPKPGQPGPPQGG 187  
 QY 117 ---EGPLPAARPGATLERAKTIS-----PGKNGVAVDVFAGAVENPEYLTPOGGAAP 169  
 DB 188 NRPQGPPEPKPGQPPPGGKSRSPQSPPEK-----PGQPPPGGNGQ 231  
 QY 170 PHPPPAFSPAFDLYWDQDPPERG-----APSTFK 201  
 DB 232 GPPPPPKP-----QGPPPGGNGKPGPPPPKPGQPPPPGSGSKSASAPAPGKQ 282

QY 202 GTPAE--NPE 210  
 DB 283 GPCQEGNNPQ 293

RESULT 11  
 ID ODO2\_MCTU STANDARD: PRT: 553 AA.  
 AC 010381:  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Dihydroliipoamide succinyltransferase component of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61) (E2).  
 GN SUCB OR RV2215 OR MT2272 OR MTCY190.26.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteriia (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock R., Basham D., Brown D., Chillingworth T., Connor R., Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."  
 RT Nature 393:537-544(1998).  
 RL  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Knouri H., Gill J., Mikula A., Bishai W.;  
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."  
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 RL  
 CC -1- FUNCTION: THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL CONVERSION OF 2-OXOGLUTARATE TO SUCCINYL-COA & CO(2). IT CONTAINS MULTIPLE COPIES OF 3 ENZYMATIC COMPONENTS: 2-OXOGLUTARATE DEHYDROGENASE (E1), DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE (E2) AND LIPOAMIDE DEHYDROGENASE (E3) (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: succinyl-CoA + dihydrolipoamide - CoA + S-succinylhydrolipoamide.  
 CC -1- COFACTOR: CONTAINS TWO COVALENTLY-BOUND LIPOYL COFACTOR (POTENTIAL).  
 CC -1- PATHWAY: tricarboxylic acid cycle.  
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL SYMMETRY (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 LIPOYL-BINDING DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: Z70283; CAA94256.1; -  
 DR EMBL: AE007072; AAK6557.1; -  
 DR HSSP: P07016; IcdT.  
 DR TIGR: MT2272; -  
 DR Tuberculist; RV2215; -

DR InterPro: IPR001078; 2oxoacid\_dh.  
 DR InterPro: IPR000089; biotin\_lipoyl.  
 DR InterPro: IPR004167; e3\_binding.  
 DR InterPro: IPR003016; lipoyl.  
 DR Pfam: PF00198; 2-oxoacid\_dh; 1.  
 DR Pfam: PF00364; biotin\_lipoyl; 2.  
 DR Pfam: PF02817; e3\_binding; 1.  
 DR ProDom: PD001115; 2oxoacid\_dh; 1.  
 DR PROSITE: PS00189; lipoyl; 2.  
 KW Tricarboxylic acid cycle; transferase; acyltransferase; lipoyl;  
 KM Complete proteome.  
 FT BINDING 43 LIPOYL (POTENTIAL).  
 FT BINDING 162 LIPOYL (POTENTIAL).  
 FT ACT\_SITE 523 BY SIMILARITY.  
 FT ACT\_SITE 527 BY SIMILARITY.  
 SO SEQUENCE 553 AA; 57087 MW; 54B6E70D23B804A7 CRC64;

Query Match 11.5%; Score 135.5; DB 1; Length 553;  
 Best Local Similarity 25.9%; Pred. No. 0.11;  
 Matches 56; Conservative 24; Mismatches 75; Indels 61; Gaps 10;

QY 27 EPSEEARSP--APSGAGSDVFDGLGMAKG-----L 61  
 DB 101 KPAPEPPVQPTSGAPAGGDAKPYLMPGLGESVTEGTVIRMLKKIGDSVGYDEPLVEYST 160  
 QY 62 QSLPHDPSPLD-----RSEDPVPLPSEND--GYVAPLCSPOPEYVNOPTVRPOP- 112  
 DB 161 DKVDTEIPSPAGVLVSTISADEDATVPVGGELARIGVADIGAAPAPAPAP--VPEPA 218  
 QY 113 PSPREGPL--PAAPAGA-----TLERAKTLSPKNGVKK--DVFA 149  
 DB 219 PTPKAPAPSPAPPAQPAAGAPVYTVLVKRLASENNIDLAGVGTGVGRIRKQDVLA 278  
 QY 150 FCGAVENPEYLTP--QCGAAPQHPPPAPSPAFNL 183  
 DB 279 AAQOKRAKAPAPAAQAAAAAPAPAPAPAPALML 314

RESULT 12  
 ID EXTN\_TOBAC STANDARD: PRT: 620 AA.  
 AC P13983;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Extensin precursor (Cell wall hydroxyproline-rich glycoprotein).  
 GN HRCPTN3.  
 OS Nicotiana tabacum (Common tobacco).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Xanthi; TISSUE=leaf;  
 RX MEDLINE=90128263; PubMed=2612909;  
 RA Keller B., Lamb C.J.;  
 RT "Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation."  
 RL Genes Dev. 3:1639-1646(1989).  
 CC -1- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE MAIN ROOT.  
 CC -1- SUBCELLULAR LOCATION: Extracellular matrix.  
 CC -1- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN GLYCOSYLATED.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.tsb-sib.ch/announce/> or send an email to [license@tsb-sib.ch](mailto:license@tsb-sib.ch)).

DR EMBL: X13885; CAA32090.1; -

DR PIR: S06733; S06733.

KW Repeat: Cell wall, Glycoprotein; Signal; Structural protein;

KM Hydroxylation.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 620 EXTENSIN.

FT REPEAT 70 73 H-A-P-P.

FT REPEAT 148 151 H-A-P-P.

FT DOMAIN 229 242 2 x 7 AA TANDEM REPEATS OF T-H-R-H-A-P-P.

FT REPEAT 229 235 1.

FT REPEAT 236 242 2.

FT DOMAIN 205 620 CONTAINS THE SER-PRO(4) REPEATS.

FT DOMAIN 499 600 3 x APPROXIMATE TANDEM REPEATS.

FT SEQUENCE 620 AA: 65406 MW, 641DD2278AB28524 CRC64;

Query Match 11.4%; Score 135; DB 1: Length 620;

Best Local Similarity 26.9%; Pred. No. 0.13;

Matches 66; Conservative 18; Mismatches 79; Indels 82; Gaps 14;

QY 18 GGSDLTGLTSESEAPRPL---APSEGAGSDVFDGDLGMAKGLQSLPTH-DPSPLQ 73

DB 130 GPCHL-----PSHGQRPSPSHGHAPSG-----GHTPRRG-QHPSPHRRSPSPS 173

QY 74 R-----YSEDPYVLPSETDGYVAPLCSQPEXVNPQDV-----RPPSPSPREG 118

DB 174 RHGHPPPTVAQPPPTIYSPSPQVQPPPTIYSPSPPTIYVQPTSPSPRGHQPPPTRHRA 233

QY 119 PLGAARAGATLE-----RAKTLSPGKNGVKNVFAAGAVENPEYITPQ 153

DB 234 P-PTNHRAPTHOPSPRLHLPSPRROROPPTYSPPRAYAQ-----SPQSPPTYSPP 286

QY 164 GGAAPQ-----HPPAFSPA-----FDNLVYWDQDPPRGAPPTSTKGTPTAENP 209

DB 287 PTYSPSPSPPTYSPPRAYSPSPPTPTPTF-----SPPPRAYSPSPPTYSPP-----P 335

QY 210 EYGL 214

DB 336 TYLPL 340

RESULT 13

EGFR\_MOUSE

ID EGFR\_MOUSE STANDARD: PRT; 1210 AA.

AC 001279;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Epidermal growth factor receptor precursor (EC 2.7.1.112).

GN EGFR.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-Liver;

RX MEDLINE-93126380; PubMed-1408137;

RA Avivi A., Skorecki K., Yaron A., Glivol D.;

RT "Promoter region of the murine fibroblast growth factor receptor 2 (bek/KGFR) gene."

RL Oncogene 7:1957-1962(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C, and CD-1; TISSUE-Liver, and Decidua;

RX MEDLINE-93126380; PubMed-7678348;

RA Paria B.C., Das S.K., Andrews G.K., Dey S.K.;

RT "Expression of the epidermal growth factor receptor gene is regulated in mouse blastocysts during delayed implantation."

RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-Liver;

RA Hibbs M.L.;

RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN-B6/C3; TISSUE-Liver;

RX MEDLINE-94170986; PubMed-8125255;

RA Luetteke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S., Jenkins N.A., Lee D.C.;

RT "The mouse waved-2 phenotype results from a point mutation in the EGF receptor tyrosine kinase."

RL Genes Dev. 8:399-413(1994).

RN [5]

RP SEQUENCE OF 1-714 FROM N.A.

RX TISSUE-Brain;

RX MEDLINE-91232866; PubMed-2030916;

RA Avivi A., Lax I., Ullrich A., Schlessinger J., Glivol D., Morse B.;

RT "Comparison of EGF receptor sequences as a guide to study the ligand binding site."

RL Oncogene 6:673-676(1991).

RN [6]

RP SEQUENCE OF 969-1117 FROM N.A.

RC STRAIN-C3H;

RA Elstinger D.P., Serrero G.;

RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF, AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein tyrosine phosphate.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to complex, internalization, internalization of the EGF-receptor complex, induction of the tyrosine kinase activity, stimulation of cell DNA synthesis, and cell proliferation.

CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.tsb-sib.ch/announce/> or send an email to [license@tsb-sib.ch](mailto:license@tsb-sib.ch)).

CC

EMBL: X78987; CAA55587.1; -

EMBL: U03425; AA17899.1; -

EMBL: X59698; CAA42219.1; -

EMBL: L06864; AAS3029.1; -

EMBL: Z12608; CAA78249.1; -

HSSP: P11362; IFGK.

MGI: 95294; Egfr.

InterPro: IPR000484; EGFR\_L\_domain.

InterPro: IPR000719; Euk\_pkinase.

InterPro: IPR002174; Furln-1like.

InterPro: IPR001245; Tyr\_pkinase.

Pfam: PF00069; pkinase; 1.

Pfam: PF00757; Furln-1like; 1.

Pfam: PF01030; Recep\_L\_domain; 2.

ProDom: PD000001; Euk\_pkinase; 1.

SMART: SM00261; FU; 3.

SMART: SM00219; TYRK; 1.

PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.

PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.

PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.

KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;

KW Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.

FT SIGNAL 1 24 POTENTIAL.

FT CHAIN 25 1210 EPIDERMAL GROWTH FACTOR RECEPTOR.

FT DOMAIN 25 647 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 648 670 POTENTIAL.

FT DOMAIN 671 1210 CYTOPLASMIC (POTENTIAL).

```

FT REPEAT 75 300 APPROXIMATE.
FT REPEAT 390 600 APPROXIMATE.
FT DOMAIN 1028 1071 SER-RICH.
FT NP_BIND 714 981 PROTEIN KINASE.
FT BINDING 720 728 ATP (BY SIMILARITY).
FT ACT_SITE 839 839 ATP (BY SIMILARITY).
FT DISULFID 190 199 BY SIMILARITY.
FT DISULFID 194 207 BY SIMILARITY.
FT DISULFID 215 223 BY SIMILARITY.
FT DISULFID 219 231 BY SIMILARITY.
FT DISULFID 232 240 BY SIMILARITY.
FT DISULFID 236 248 BY SIMILARITY.
FT DISULFID 251 260 BY SIMILARITY.
FT DISULFID 264 291 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 311 326 BY SIMILARITY.
FT DISULFID 329 333 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 526 535 BY SIMILARITY.
FT DISULFID 539 555 BY SIMILARITY.
FT DISULFID 558 571 BY SIMILARITY.
FT DISULFID 562 579 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT DISULFID 595 617 BY SIMILARITY.
FT DISULFID 620 628 BY SIMILARITY.
FT DISULFID 624 636 BY SIMILARITY.
FT MOD_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-, MAJOR SITE) (BY SIMILARITY).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 19 19 C -> S (IN REF. 2).
FT CONFLICT 539 539 C -> W (IN REF. 5).
FT CONFLICT 991 991 L -> F (IN REF. 4).
FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DF2D2F5 CRC64;

```

Query Match 11.38; Score 134; DB 1; Length 1210;  
 Best Local Similarity 26.48; Pred. No. 0.27; Mismatches 63; Indels 76; Gaps 11;  
 Matches 58; Conservative 23;

```

QY 47 DVEDGDLGMAKGLQSLPTHPSP-----LQRYYS 76
  ||| : : : : :
DB 1011 DVVDADAEYLTPQGEFFNSPTSLSLATSNNSTVACINNGSCRYKEDAFLOQYS 1070
  ||| : : : : :
QY 77 EDPIVPLPSEF--DGYVAPLTCSPQPEYVNPQVRRPSPREGLPLAARAGATLEPAK 134
  ||| : : : : :
DB 1071 SDPIGAVTEENIDAFI-----PVPEYVNO-SYPKRRAGSVQNPVYHNOPLRP----- 1117
  ||| : : : : :
QY 135 TLSPGKGVKVDYFAFGAVENPEYL--TPQGAAPQPHRPAFPAPDNLYWDO----- 188
  ||| : : : : :
DB 1118 --APGRLLHON--PHSNAVGNPEYLNTAQ-----PCLSSGGRSPALMILQKSHQ 1164
  ||| : : : : :
QY 189 -----DP-----PBGAPSTFGKPTAENPEYLGDPV 217
  ||| : : : : :
DB 1165 MSLDNPDYQODFFPKETKPNICFKG-PTAENAEYLRVAP 1203
  ||| : : : : :

```

RESULT 14  
 SM6B\_MOUSE

```

ID SM6B_MOUSE STANDARD; PRT; 886 AA.
AC 054951:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 6B precursor (Semaphorin VIB) (Sema VIB) (Semaphorin N) (Sema N).
GN SEMA6B OR SEMAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98027184; PubMed=9361278;
RA Eckhardt F., Behar O., Calautti E., Yonezawa K., Nishimoto I.,
RA Fishman M.C.;
RT "A novel transmembrane semaphorin can bind c-src.";
RL Mol. Cell. Neurosci. 9:409-419(1997).
CC -1- SUBUNIT: HOMODIMER. BINDS SPECIFICALLY THE SH3 DOMAIN OF THE PROTOONCOGENE C-SRC.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: DURING DEVELOPMENT IT IS EXPRESSED IN MUSCLE. IN ADULTHOOD, IT IS EXPRESSED UBICUOUSLY.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF036585; AAC00493.1;
CC MGD: MGI:1202889; Sema6b.
CC Interpro: IPR003659; Plexin-like.
CC Interpro: IPR001627; Sema.
CC DR Pfam: PF01403; Sema; 1.
CC SMART: SM00423; PSI; 1.
CC KMW Signal: Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
CC KM developmental protein.
FT FT CHAIN 1 26 POTENTIAL.
FT FT SIGNAL 27 886 SEMAPHORIN 6B.
FT FT DOMAIN 27 605 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 606 626 POTENTIAL.
FT FT DOMAIN 627 886 CYTOPLASMIC (POTENTIAL).
FT FT DOMAIN 239 549 SEMA.
FT FT DOMAIN 751 754 POLY-LEU.
FT FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 886 AA; 95466 MW; E5F56D125CDA57AD CRC64;

```

Query Match 11.08; Score 130.5; DB 1; Length 886;  
 Best Local Similarity 25.88; Pred. No. 0.33; Mismatches 93; Indels 59; Gaps 10;  
 Matches 61; Conservative 23;

```

QY 1 GAGG-----VWHRHRSSTRSGGDLTGLTEPSEEPAP-----RSPLAPS 41
  ||| : : : : :
DB 669 GAGGPGGPPALLAPLQNKQMTAAALHGSPHDLDTGLTPDTPTLPQRLPPIPHNA 728
  ||| : : : : :
QY 42 EGAGSDVFDGDLGMAKGLQSLPTHPSPPLQRYSEDPVPL-PSETDGYVAPLTC----- 96
  ||| : : : : :
DB 729 HALGSRAMDHSHALLSASASILLAPV---RASEQGVYAPAEPPESRLCAPSCRAHS 785
  ||| : : : : :
QY 97 -----SPOPEYVNPQVRRPQPSPRPG-PLPAARPAAGATLERAKTLSPGKNG--- 142
  ||| : : : : :

```

Db 786 PGDFPLTPHASPSPRRVVSAPTGPDLPSVGDGLPGWSPSPATSSLR-----PGPHGPP 840  
QY 143 -VVKDVFAGGAVENPEYLTPGGAAPQHPHPPAFSPAFDNLXYWDQDPPERGA 197  
Db 841 AALRRHTHTFNGSEARP-----GCHRRPRRHPA-----DSTHLPLCGCTERTAPP 884

## RESULT 15

PRP2\_HUMAN STANDARD; PRT; 251 AA.  
ID PRP2\_HUMAN  
AC P02812;  
DT 21-JUL-1986 (rel. 01, Created)  
DT 13-AUG-1987 (rel. 05, Last sequence update)  
DT 16-OCT-2001 (rel. 40, Last annotation update)  
DE Salivary proline-rich protein precursor (Clone CP7) [Contains: Basic  
DE peptide P-F] (Fragment).  
GN PRB2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85289325; PubMed=2993301;  
RA Maeda N., Kim H.-S., Azen E.A., Smithies O.;  
RT "Differential RNA splicing and post-translational cleavages in the  
RT human salivary proline-rich protein gene system.";  
RL J. Biol. Chem. 260:11123-11130(1985).  
RN (2)  
RP SEQUENCE OF 134-194.  
RX MEDLINE=83265674; PubMed=6874669;  
RA Salton E., Isemura S., Sanada K.;  
RT "Complete amino acid sequence of a basic proline-rich peptide, P-F,  
RT from human parotid saliva.";  
RL J. Biochem. 93:883-888(1983).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; K03208; AAA60189.1; -  
DR PIR; A03294; PIHUPF.  
DR PIR; E25372; E25372.  
DR Genew; HGNC:9338; PRB2.  
DR MIM; 168810; -  
KW Repeat; Parotid gland; Multigene family; Saliva.  
FT NON\_TER 1 1  
FT CHAIN 134 194 BASIC PEPTIDE P-F.  
FT REPEAT <1 9  
FT REPEAT 10 71  
FT REPEAT 72 133  
FT REPEAT 134 195  
SQ SEQUENCE 251 AA; 24641 MW; D779F590C0EBF30B CRC64;  
Query Match 11.0%; Score 129.5; DB 1; Length 251;  
Best local similarity 25.1%; Pred. No. 0.11;  
Matches 55; Conservative 15; Mismatches 88; Indels 61; Gaps 9;  
QY 10 HRSSTRSGGGDLTLGLEPSEEAAPS-PLAPSEAGSDVFDGLGKAAGLQSLPTHD 68  
Db 3 NKSRRSSRPPGK-PPGPPPGGNGPQGGPPPGKPPGCGNKPPGPP--PGKPGCP 59  
QY 69 PSLPQRTSEDPYVPLSENDGYAPLTCSPQPRYVNPQVPRPQPPSPREGPLP----- 121  
Db 60 PPGDNKSGSARSPP-PPGPG-----PPPGGNGPQGGPPPPPGKPPGPPPGGDNKSG 110  
QY 122 -----AARPAGATLEAKTLSPKNGVAVDFAGGAVENPEYLTPGGAAPQHP 172  
Db 111 GPPPPGKPPGPPPGGSKSSRSRPPGK-----PPGPPPGGNGPQGGPP 154

QY 173 PPAFSPAFDNLXYWDQDPPERGA-----PSTFKGTP 204  
Db 155 PPGKP-----OGPPPGGNGKPPGPPPPGKPPGPP 184

Search completed: April 28, 2003, 13:40:16  
Job time : 8.53595 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:23 ; Search time 6.04097 Seconds  
(without alignments)  
1984.228 Million cell updates/sec

Title: US-09-821-883-23  
Perfect score: 1587

Sequence: 1 STGYCTGDMKRLRPASPET.....FGASCYACPFYNTSDVGS 289

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1587	100.0	1255	1	ERRB2_HUMAN
2	1353.5	85.3	1257	1	ERRB2_RAT
3	1322	83.3	1254	1	ERRB2_MESAU
4	739	46.6	703	1	EGFR_CHICK
5	716.5	45.1	1210	1	EGFR_HUMAN
6	716.5	45.1	1210	1	EGFR_MOUSE
7	716.5	45.1	1308	1	ERRB4_HUMAN
8	706.5	44.5	1308	1	ERRB4_RAT
9	686	43.2	1342	1	ERRB3_HUMAN
10	658	41.5	1339	1	ERRB3_RAT
11	640.5	40.4	1167	1	XRMRK_XIPMA
12	535	33.7	1426	1	EGFR_DROME
13	396.5	25.0	1323	1	LT23_CAEEL
14	315.5	19.9	1363	1	ILPR_BRALA
15	273.5	17.2	1477	1	HTK7_HYDAT
16	255	16.1	2146	1	INSR_DROME
17	246.5	15.5	1300	1	IRR_CAVPO
18	246.5	15.5	1607	1	MRPR_LYMAST
19	245.5	15.5	1300	1	IRR_MOUSE
20	244.5	15.4	1382	1	INSR_HUMAN
21	244.5	15.4	1383	1	INSR_RAT
22	243.5	15.3	1372	1	INSR_MOUSE
23	238.5	15.0	1297	1	IRR_HUMAN
24	232.5	14.7	1367	1	IGIR_HUMAN
25	232.5	14.7	1370	1	IGIR_RAT
26	230.5	14.5	1373	1	IGIR_MOUSE
27	228.5	14.4	581	1	IRR_RAT
28	207	13.0	1390	1	INSR_AEDAE
29	140	8.8	1696	1	PKC5_BRACL
30	125	7.9	1875	1	PKC5_MOUSE
31	121	7.6	1680	1	FUR2_DROME
32	108.5	6.8	798	1	ITB0_XENLA
33	106	6.7	1877	1	PKC5_RAT

34	105.5	6.6	798	1	ITB1_XENLA
35	105	6.6	830	1	SREC_HUMAN
36	104	6.6	937	1	PAC4_RAT
37	103.5	6.5	769	1	ITB2_HUMAN
38	101.5	6.4	484	1	LEM2_PIG
39	101.5	6.4	969	1	PAC4_HUMAN
40	100	6.3	551	1	LEM2_RABIT
41	98	6.2	3075	1	LMN1_HUMAN
42	97	6.1	837	1	ATSA_HUMAN
43	97	6.1	3106	1	LMN2_MOUSE
44	96.5	6.1	2282	1	ZAN_RABIT
45	96	6.0	667	1	TS11_GIALA

## ALIGNMENTS

RESULT 1	ID	ERRB2_HUMAN	STANDARD:	PRT: 1255 AA.
AC	P04626:			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DE	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC:2.7.1.112)			
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell			
DE	surface receptor HER2) (MIM 19).			
CN	ERRB2 OR HER2 OR NGL OR NEU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86118663; PubMed=3003577;			
RA	Yamamoto T., Ikawa S., Akiyama T., Samba K., Nomura N., Miyajima N.,			
RA	Saito T., Toyoshima K.;			
RT	"Similarity of protein encoded by the human c-erbB-2 gene to			
RT	epidermal growth factor receptor.";			
RL	Nature 319:230-234(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86070181; PubMed=2999974;			
RA	Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,			
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger T.,			
RT	Francine U., Levinson A., Ullrich A.;			
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor			
RT	shares chromosomal location with neu oncogene.";			
RL	Science 230:1132-1139(1985).			
RN	[3]			
RP	SEQUENCE OF 737-1031 FROM N.A.			
RX	MEDLINE=86016729; PubMed=2995967;			
RA	Samba K., Kanata N., Toyoshima K., Yamamoto T.;			
RA	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the			
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a			
RT	human salivary gland adenocarcinoma.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).			
RN	[4]			
RP	VARIANTS VAL-654 AND VAL-655.			
RX	MEDLINE=93194196; PubMed=8095488;			
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;			
RT	"Characterization of a new allele of the human ERBB2 gene by allele-			
RT	specific competition hybridization.";			
RL	Genomics 15:426-429(1993).			
CC	-1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,			
CC	ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A			
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-			
CC	ALPHA AND AMBIREGULIN.			
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein			
CC	tyrosine phosphate.			
CC	-1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS			
CC	(POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			

CC -1- P-TM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN  
 CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY  
 CC OF 0.783; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;  
 CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M11767: AAA35808.1: -  
 DR EMBL: M11761: AAA35808.1: JOINED.  
 DR EMBL: M11763: AAA35808.1: JOINED.  
 DR EMBL: M11763: AAA35808.1: JOINED.  
 DR EMBL: M11764: AAA35808.1: JOINED.  
 DR EMBL: M11765: AAA35808.1: JOINED.  
 DR EMBL: M11766: AAA35808.1: JOINED.  
 DR EMBL: M11730: AAA75493.1: -  
 DR EMBL: M12036: AAA35978.1: -  
 DR EMBL: X03363: CAA27060.1: -  
 DR PIR: A25491: A25491.  
 DR PIR: A24571: A24571.  
 DR HSSP: P11362: IFGK.  
 DR Genew: HGNC:3430; ERBB2.  
 DR MIM: 164870: -  
 DR InterPro: IPR000494: EGFR\_L\_domain.  
 DR InterPro: IPR000719: Euk\_pkinase.  
 DR InterPro: IPR002174: Furin-like.  
 DR InterPro: IPR001245: Tyr\_pkinase.  
 DR InterPro: IPR004019: YLP\_molif.  
 DR Pfam: PF00069: pkinase.1.  
 DR Pfam: PF00757: Furin-like.1.  
 DR Pfam: PF01030: Recep\_L\_domain.2.  
 DR Pfam: PF02757: YLP.2.  
 DR ProDom: PD000001: Euk\_pkinase.1.  
 DR SMART: SM00261: FU.3.  
 DR SMART: SM00219: TYRKC.1.  
 DR PROSITE: PS00107: PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00109: PROTEIN\_KINASE\_TYR.1.  
 DR PROSITE: PS50011: PROTEIN\_KINASE\_DOM.1.  
 KM Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;  
 KM Transferrase: Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KM Polymorphism.  
 FT SIGNAL 1 21  
 FT CHAIN 22 1255  
 FT DOMAIN 22 652  
 FT TRANSMEM 653 675  
 FT DOMAIN 676 1255  
 FT DOMAIN 720 987  
 FT NP\_BIND 726 734  
 FT BINDING 753 753  
 FT ACT\_SITE 845 845  
 FT DISULFID 199 204  
 FT DISULFID 220 227  
 FT DISULFID 224 225  
 FT DISULFID 236 244  
 FT DISULFID 240 252  
 FT DISULFID 255 264  
 FT DISULFID 268 295  
 FT DISULFID 299 311  
 FT DISULFID 315 331  
 FT DISULFID 334 338  
 FT DISULFID 511 520  
 FT DISULFID 515 528  
 FT DISULFID 531 540  
 FT DISULFID 544 560  
 FT DISULFID 560

FT DISULFID 563 576 BY SIMILARITY.  
 FT DISULFID 567 584 BY SIMILARITY.  
 FT DISULFID 587 596 BY SIMILARITY.  
 FT DISULFID 600 623 BY SIMILARITY.  
 FT DISULFID 626 634 BY SIMILARITY.  
 FT DISULFID 630 642 BY SIMILARITY.  
 FT MOD\_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 124 124 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 530 530 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 571 571 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 629 629 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT VARIANT 654 654 I -> V.  
 FT VARIANT 654 654 /FTID-VAR\_004077.  
 FT VARIANT 655 655 I -> V.  
 FT VARIANT 655 655 /FTID-VAR\_004078.  
 FT CONFLICT 1170 1170 P -> A (IN REF. 2).  
 FT SEQUENCE 1255 AA: 137909 MW: 39E9DFDA04DCF962 CRC64:  
 SO  
 Query Match 100.0%; Score 1587; DB 1; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 9, 2e-127; Indels 0; Gaps 0;  
 Matches 289; Conservative 0; Mismatches 0;  
 QY 1 STGYCTGDMKRLRLPASPETHLMLRLHLYGCGVQGNLELTYLPTNASTSFLQDIOEQV 60  
 DB 22 STGYCTGDMKRLRLPASPETHLMLRLHLYGCGVQGNLELTYLPTNASTSFLQDIOEQV 81  
 QY 61 GYVLIAHNOVROVPLQRLIRVGTQLEFDNYALAVLDNGPLNNTPTVTASRGGLRELQ 120  
 DB 82 GYVLIAHNOVROVPLQRLIRVGTQLEFDNYALAVLDNGPLNNTPTVTASRGGLRELQ 141  
 QY 121 LRSITLTKGCVLQIRNPOLCYOPTILMKDIFPKNNOLATLIDTNSRACHCSPCKG 180  
 DB 142 LRSITLTKGCVLQIRNPOLCYOPTILMKDIFPKNNOLATLIDTNSRACHCSPCKG 201  
 QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHEBOCAAGCTGPKHSCLACLHNHS 240  
 DB 202 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHEBOCAAGCTGPKHSCLACLHNHS 261  
 QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYISTDVG 289  
 DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYISTDVG 310  
 RESULT 2  
 ERB2\_RAT  
 ID ERB2\_RAT STANDARD: PRT: 1257 AA.  
 AC P06494;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor  
 DE receptor-related protein).  
 GN ERBB2 OR NEU.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN (1)  
 RC SEQUENCE FROM N.A.  
 RC TISSUE-Neuroblastoma;  
 RX MEDLINE=86118662; PubMed=3945311;  
 RA Bergmann C.I., Hung M.-C., Weinberg R.A.;  
 RT "The neu oncogene encodes an epidermal growth factor receptor-related  
 RL protein.";  
 RL Nature 319:226-230(1986).  
 RP SEQUENCE OF 852-905 FROM N.A.  
 RC TISSUE-Sciatic nerve;

RX MEDLINE-91222560; PubMed-2025425;  
 RA Lai C., Lemke G.;  
 RT "An extended family of protein-tyrosine kinase genes differentially  
 RT expressed in the vertebrate nervous system.";  
 RL Neuron 6:691-704(1991).  
 RN [3]  
 RP STRUCTURE BY NMR OF 650-668.  
 RX MEDLINE-92155181; PubMed-1346763;  
 RA Gullick W.J., Bottomley A.C., Lofts F.J., Doak D.G., Mulvey D.,  
 RA Newman R., Crumpton M.J., Sternberg M.O.E., Campbell I.D.;  
 RT "Three dimensional structure of the transmembrane region of the proto-  
 RT oncogenic and oncogenic forms of the neu protein.";  
 RL EMBO J. 11:43-48(1992).  
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX.  
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A  
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-  
 CC ALPHA AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS.  
 CC THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PFM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X03362; CAA27059.1; ALT\_INIT.  
 DR PIR: A24562; TVRTNU.  
 DR HSP: P11362; IEGK.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Purln-like.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR InterPro: IPR004019; YLP\_molif.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR Pfam: PF00757; Furln-like; 1.  
 DR Pfam: PF02030; Recep\_L\_domain; 2.  
 DR Pfam: PF02757; YLP; 2.  
 DR ProDom: PD000001; Euk\_Pkinase; 1.  
 DR SMART: SM00261; FU; 3.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Proto-oncogene; Disease mutation.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 1257 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.  
 FT DOMAIN 22 654 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 655 677 POTENTIAL.  
 FT DOMAIN 678 1257 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 159 369 CYS-RICH.  
 FT DOMAIN 473 646 CYS-RICH.  
 FT DOMAIN 722 989 PROTEIN KINASE.  
 FT NP\_BIND 728 736 ATP (BY SIMILARITY).  
 FT BINDING 755 755 ATP (BY SIMILARITY).  
 FT ACT\_SITE 847 847 BY SIMILARITY.  
 FT DISULFID 196 205 BY SIMILARITY.  
 FT DISULFID 200 213 BY SIMILARITY.  
 FT DISULFID 221 228 BY SIMILARITY.  
 FT DISULFID 225 236 BY SIMILARITY.  
 FT DISULFID 237 245 BY SIMILARITY.  
 FT DISULFID 241 253 BY SIMILARITY.  
 FT DISULFID 256 265 BY SIMILARITY.

FT DISULFID 269 296 BY SIMILARITY.  
 FT DISULFID 300 312 BY SIMILARITY.  
 FT DISULFID 316 332 BY SIMILARITY.  
 FT DISULFID 335 339 BY SIMILARITY.  
 FT DISULFID 513 522 BY SIMILARITY.  
 FT DISULFID 517 530 BY SIMILARITY.  
 FT DISULFID 533 542 BY SIMILARITY.  
 FT DISULFID 546 562 BY SIMILARITY.  
 FT DISULFID 565 578 BY SIMILARITY.  
 FT DISULFID 569 586 BY SIMILARITY.  
 FT DISULFID 589 598 BY SIMILARITY.  
 FT DISULFID 602 625 BY SIMILARITY.  
 FT DISULFID 628 636 BY SIMILARITY.  
 FT DISULFID 632 644 BY SIMILARITY.  
 FT MOD\_RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 188 188 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 532 532 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 573 573 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 631 631 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).  
 SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64.  
 Query Match 85.3%; Score 1353.5; DB 1; Length 1257;  
 Best Local Similarity 85.5%; Pred. No. 5.7e-107;  
 Matches 247; Conservative 13; Mismatches 28; Indels 1; Gaps 1;  
 QY 2 TQVCTGDMKRLRLPASPETHIDMLRHLYGCGQVVGNNLELYPTNMSLSFLADTQIEVGG 61  
 DB 23 TQVCTGDMKRLRLPASPETHIDMLRHLYGCGQVVGNNLELYPTNMSLSFLADTQIEVGG 82  
 QY 62 YVLAHNVGRVPRQRLRIYRGTOLEFEDNVLAVIDNDPLNNTPTV-GASPGGLREIQ 120  
 DB 83 YVLAHNVGRVPRQRLRIYRGTOLEFEDNVLAVIDNDPLNNTPTV-GASPGGLREIQ 142  
 QY 121 LRLSTLEILKGVLLQIQRNPOLCYQDPIILMKDIFHKKNNQLATLIDITNNSRACHPCSPMCKG 180  
 DB 143 LRLSTLEILKGVLLQIQRNPOLCYQDPIILMKDIFHKKNNQLAPVDIDITNNSRACHPCSPMCKG 202  
 QY 181 SRCWGESSEDCQSLTRYVACGACGACRCKGRLPTDCCHQCAAGCTGPKHSDCLAEHFNHS 240  
 DB 203 NHCWGESSEDCQSLTRYVACGACGACRCKGRLPTDCCHQCAAGCTGPKHSDCLAEHFNHS 262  
 QY 241 GICELHCPALVTYNTDFEESHPNPEGRTFGASCVTACPYVLTSDVCS 289  
 DB 263 GICELHCPALVTYNTDFEESHPNPEGRTFGASCVTACPYVLTSDVCS 311  
 RESULT 3  
 ERB2\_MESAU STANDARD; PRT: 1254 AA.  
 ID ERB2\_MESAU  
 AC Q60553;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)  
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).  
 GN ERBB2 OR NEU.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 CC Mesocricetus.  
 OX NCBI\_TaxId=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Nerve;  
 RX MEDLINE-94193007; PubMed-7908275;  
 RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,  
 RA Yamazaki Y., Ishikawa T.;  
 RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";  
 RL Gene 140:251-255(1994).



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: M20386; AAA48760.1; -  
 DR InterPro: IPR000494; EGFR\_Ldomain.  
 DR InterPro: IPR000719; Euk\_Pkinase.  
 DR InterPro: IPR002174; Furin-Like.  
 DR InterPro: IPR001245; Tyr\_Pkinase.  
 DR Pfam: PF00757; Furin-Like; 1.  
 DR Pfam: PF01030; Recep\_Ldomain; 2.  
 DR SMART: SMO0261; FU: 4.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP\_PARTIAL.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR\_PARTIAL.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM\_PARTIAL.  
 DR Transmembrane: Glycoprotein; Receptor; Signal; transferase;  
 KW Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
 FT SIGNAL 1 30  
 FT CHAIN 1 >703 EPIDERMAL GROWTH FACTOR RECEPTOR.  
 FT DOMAIN 31 654 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 655 667 POTENTIAL.  
 FT DOMAIN 668 >703 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 197 206 BY SIMILARITY.  
 FT DISULFID 201 214 BY SIMILARITY.  
 FT DISULFID 222 230 BY SIMILARITY.  
 FT DISULFID 226 238 BY SIMILARITY.  
 FT DISULFID 239 247 BY SIMILARITY.  
 FT DISULFID 243 255 BY SIMILARITY.  
 FT DISULFID 258 267 BY SIMILARITY.  
 FT DISULFID 271 298 BY SIMILARITY.  
 FT DISULFID 302 314 BY SIMILARITY.  
 FT DISULFID 318 333 BY SIMILARITY.  
 FT DISULFID 336 340 BY SIMILARITY.  
 FT DISULFID 513 522 BY SIMILARITY.  
 FT DISULFID 517 530 BY SIMILARITY.  
 FT DISULFID 533 542 BY SIMILARITY.  
 FT DISULFID 546 562 BY SIMILARITY.  
 FT DISULFID 565 581 BY SIMILARITY.  
 FT DISULFID 589 598 BY SIMILARITY.  
 FT DISULFID 592 601 BY SIMILARITY.  
 FT DISULFID 605 627 BY SIMILARITY.  
 FT DISULFID 630 638 BY SIMILARITY.  
 FT DISULFID 634 646 BY SIMILARITY.  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 578 578 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT NON\_TER 703 703  
 SQ SEQUENCE 703 AA; 77427 MW; AEF2DE1B735A690 CRC64;

Query Match 46.6%; Score 739; DB 1; Length 703;  
 Best Local Similarity 48.1%; Pred. No. 3.6e-55;  
 Matches 139; Conservative 48; Mismatches 90; Indels 12; Gaps 5;

OY 3 OVCSTGDMKRLPASPETHLDMLRLHLYGCGQVYVGNLELTYLPTNASLSFLQDIDQVQY 62  
 DB 35 KVCQGTNNKLTQLGHEVDHFTSLQRMVNNCEVVLNSLETTYEHNRDLFLTKIQEVAGY 94  
 OY 63 VLIANNQROYVLOLRIRYRGQLFEDYNALAVLNDGNPLNTTIVTGSFGGLKQLDR 122  
 DB 95 VLIANNQVYVLOLRIRYRGQLFEDYNALAVLNDGNPLNTTIVTGSFGGLKQLDR 145  
 OY 123 SLTEILKGVGLIQRNPQLCYQDTILMKDIFHNKQDLATLID-TNRSRACHPCSPWCKGS 181

DB 146 RLSEILNGVATISNNPKLCNNDYLMNDIDITSRK-PLTVLDFASNLSSCPKCPNCTED 204  
 OY 182 RCGWESSDCCSLRTYVACGCA-RCKGPLETDCHECCAGCGPKHSD-LACLHFNHS 240  
 DB 205 HCGWGEQNCOTLTKVICAQCCSGCRGKVPDSCCHNCAAGCGPSPSDIACRFRDD 264  
 OY 241 GICELHCPALVTYNTDFESPMPNPGRTFGASCCTACPYNYLSTVGS 289  
 DB 265 ATCKDTCPLVLYNFTTYQMDVNEBKISFGATCYRBCPHNYVTIDHS 313

## RESULT 5

EGFR\_HUMAN STANDARD; PRT: 1210 AA.  
 ID EGFR\_HUMAN  
 AC P00533; P06288; Q14225; Q9UMD7; Q9UMD8; Q9UMG5; Q92795; O00732;  
 AC Q00688; Q9B252; Q9H2C9; Q9G2X1; Q9H3C9;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor  
 protein-tyrosine kinase ErbB-1).  
 GN EGFR OR ERBB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_Taxid:9606;  
 RN [1]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE-84219729; PubMed-6328312;  
 RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
 RA Lee J., Yarden Y., Liberman T.A., Schlessinger J., Downward J.,  
 RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;  
 RT "Human epidermal growth factor receptor cDNA sequence and aberrant  
 RT expression of the amplified gene in A431 epidermoid carcinoma cells."  
 RL Nature 309:418-425(1984).  
 RN [2]  
 RN SEQUENCE FROM N.A. (ISOFORM 2).  
 RX TISSUE-Placenta;  
 RC MEDLINE-95382957; PubMed-7654368;  
 RA Ilekis J.V., Stark B.C., Scoccia B.;  
 RT "Possible role of variant RNA transcripts in the regulation of  
 RT epidermal growth factor receptor expression in human placenta."  
 RL Mol. Reprod. Dev. 41:149-156(1995).  
 RN [3]  
 RN SEQUENCE FROM N.A. (ISOFORM 2).  
 RX TISSUE-Placenta;  
 RC MEDLINE-97078686; PubMed-8918811;  
 RA Reiter J.L., Mahle N.J.;  
 RT "A 1.8 kb alternative transcript from the human epidermal growth  
 RT factor receptor gene encodes a truncated form of the receptor."  
 RL Nucleic Acids Res. 24:4050-4056(1996).  
 RN [4]  
 RN SEQUENCE FROM N.A. (ISOFORM 2).  
 RX TISSUE-Placenta;  
 RC MEDLINE-97256547; PubMed-9103388;  
 RA Ilekis J.V., Gattli J., Niederberger C., Scoccia B.;  
 RT "Expression of a truncated epidermal growth factor receptor-like  
 RT protein (TEGFR) in ovarian cancer."  
 RL Gynecol. Oncol. 65:36-41(1997).  
 RN [5]  
 RN SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
 RX TISSUE-Placenta;  
 RC MEDLINE-21100872; PubMed-11161793;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramanian S., Crossley T.D., Magnuson T.R., James C.D.,  
 RA Mahle N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms."  
 RL Genomics 71:1-20(2001).  
 RN [6]

RP SEQUENCE OF 575-687 FROM N.A.  
RA Reiter J.L., Thredgill D.W., Danielsen A.J., Schehl C.M.,  
RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
RA Mehlle N.J.;  
RT "human and mouse alternative EGFR transcripts encoding only the  
RT extracellular domain of the receptor.";  
RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE OF 713-924 FROM N.A.  
RX MEDLINE-84196372; PubMed-6326261;  
RA Lin C.R., Chen W.S., Krueger W., Stolarsky L.S., Weber W.,  
RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;  
RT "Expression cloning of human EGF receptor complementary DNA: gene  
RT amplification and three related messenger RNA products in A431  
RT cells.";  
RL Science 224:843-848(1984).  
RN [8]  
RP SEQUENCE OF 150-962 FROM N.A.  
RX MEDLINE-84245835; PubMed-6330563;  
RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
RA Roe B.A., Merlino G.T., Pastan I.;  
RT "Human epidermal growth factor receptor cDNA is homologous to a  
RT variety of RNAs overproduced in A431 carcinoma cells.";  
RL Nature 309:806-810(1984).  
RN [9]  
RP SEQUENCE OF 1028-1210 FROM N.A.  
RX MEDLINE-85046483; PubMed-6093780;  
RA Simmen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,  
RA O'Malley B.W.;  
RT "Isolation of an evolutionarily conserved epidermal growth factor  
RT receptor cDNA from human A431 carcinoma cells";  
RL Biochem. Biophys. Res. Commun. 124:125-132(1984).  
RN [10]  
RP SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE-88217333; PubMed-3329716;  
RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
RA Waterfield M.D.;  
RT "The human EGF receptor gene: structure of the 110 kb locus and  
RT identification of sequences regulating its transcription.";  
RL Oncogene Res. 1:375-396(1987).  
RN [11]  
RP SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE-91107677; PubMed-1988448;  
RA Haley J.D., Waterfield M.D.;  
RT "Contributory effects of de novo transcription and premature  
RT transcript termination in the regulation of human epidermal growth  
RT factor receptor proto-oncogene RNA synthesis.";  
RL J. Biol. Chem. 266:1746-1753(1991).  
RN [12]  
RP SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE-85270438; PubMed-22991899;  
RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;  
RT "Characterization and sequence of the promoter region of the human  
RT epidermal growth factor receptor gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
RN [13]  
RP SEQUENCE OF 540.  
RA Konda D.;  
RT Submitted (SEP-1997) to the SWISS-PROT data bank.  
RN [14]  
RP RECEPTOR ACTIVITY.  
RX MEDLINE-84191554; PubMed-6325948;  
RA Mroczkowski B., Mosig G., Cohen S.;  
RT "ATP-stimulated interaction between epidermal growth factor receptor  
RT and supercoiled DNA.";  
RL Nature 309:270-273(1984).  
RN [15]  
RP PHOSPHORYLATION.  
RX MEDLINE-89278137; PubMed-2543678;  
RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,  
RA Hawk R., Glyvol D., Ullrich A., Schlessinger J.;  
RT "All autophosphorylation sites of epidermal growth factor (EGF)  
RT receptor and HBR2/neu are located in their carboxyl-terminal tails.

RT Identification of a novel site in EGF receptor.";  
RL J. Biol. Chem. 264:10667-10671(1989).  
RN [16]  
RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
RX ASN-528  
RX MEDLINE-96398132; PubMed-8962717;  
RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;  
RT "Analysis of the glycosylation patterns of the extracellular domain of  
RT the epidermal growth factor receptor expressed in Chinese hamster  
RT ovary fibroblasts.";  
RL Growth Factors 13:121-132(1996).  
RN [17]  
RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND  
RX ASN-603.  
RX MEDLINE-20198209; PubMed-10731668;  
RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;  
RT "Characterization of the N-oligosaccharides attached to the atypical  
RT Asn-X-Cys sequence of recombinant human epidermal growth factor  
RT receptor.";  
RL J. Biochem. 127:65-72(2000).  
RN [18]  
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.  
RX MEDLINE-98225196; PubMed-9556602;  
RA Abe Y., Odaka M., Inagaki F., Lax I., Schlessinger J., Kohda D.;  
RT "Disulfide bond structure of human epidermal growth factor receptor.";  
RL J. Biol. Chem. 273:11150-11157(1998).  
RN [19]  
RP REVIEW.  
RX MEDLINE-87297456; PubMed-3039909;  
RA Carpenter G.;  
RT "Receptors for epidermal growth factor and other polypeptide  
RT mitogens.";  
RL Annu. Rev. Biochem. 56:881-914(1987).  
CC -1- FUNCTION: Receptor for EGF, but also for other members of the EGF  
CC family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding  
CC EGF-like growth factor, GP30 and vaccine virus growth factor. Is  
CC involved in the control of cell growth and differentiation.  
CC -1- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is  
CC secreted.  
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms, 1/p170 (shown here), 2/p60/  
CC truncated isoform/TGFR, 3/p110 and 4; are produced by  
CC alternative splicing.  
CC -1- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also  
CC expressed in ovarian cancers.  
CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to  
CC dimerization, internalization of the EGF-receptor complex,  
CC induction of the tyrosine kinase activity, stimulation of cell DNA  
CC synthesis, and cell proliferation.  
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X00588; CAA25240.1; -  
DR EMBL: U95089; AAB53063.1; -  
DR EMBL: U48722; AAC50802.1; -  
DR EMBL: U48723; AAC50804.1; -  
DR EMBL: U48724; AAC50796.1; -  
DR EMBL: U48725; AAC50797.1; -  
DR EMBL: U48726; AAC50798.1; -

Query Match 45.1%; Score 716.5; DB 1; Length 1210;  
Best Local Similarity 46.5%; Pred. No. 5.4e-53;  
Matches 134; Conservative 44; Mismatches 99; Indels 11; Gaps 2;



FT DISULFID 558 571 BY SIMILARITY.  
 FT DISULFID 562 579 BY SIMILARITY.  
 FT DISULFID 582 591 BY SIMILARITY.  
 FT DISULFID 595 617 BY SIMILARITY.  
 FT DISULFID 620 628 BY SIMILARITY.  
 FT DISULFID 624 636 BY SIMILARITY.  
 FT MOD\_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1197 1197 PHOSPHORYLATION (AUTO-, MAJOR SITE) (BY SIMILARITY).  
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 528 528 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 19 19 C -> S (IN REF. 2).  
 FT CONFLICT 539 539 C -> W (IN REF. 5).  
 FT CONFLICT 991 991 L -> F (IN REF. 4).  
 FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).  
 SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DF2D2F5 CRC64;

Query Match 45.1%; Score 716.5; DB 1; Length 1210;  
 Best local Similarity 47.3%; Pred. No. 5.4e-53;  
 Matches 138; Conservative 39; Mismatches 96; Indels 19; Gaps 4;

QY 3 QVCTGDMKRLRASPETHLDMRLHYGCGVQVQGNLELYTPMNSLSTLQDIQVQY 62  
 DB 29 KVOGGSNRLTQJGTEFHLFLSLQRMVNNCEVVLGNLEITYVORNDLSLKTQIEVAGY 88  
 QY 63 VLLAHNOVROVFORLRIYVGTQLEFNDVALVANDGDPINNTPTVYGASPGGIREQLR 122  
 DB 89 VLLAHNTVEIPLENLDIINGNALITYVALILSN-----YGRTRGRLPLMR 138  
 QY 123 SLTEILKGVLIQRPOLCYQDTITLMDI---FHKNQALTLIDTNRSRACHPCSPMC 178  
 DB 139 NLOEILILGAVRFNSNPILTCNMDTIOWRDYQVFMNSMDL---QSHRSPKOPSC 194  
 QY 179 KSGRCMGESEDDOSLTRYCAGGCA-RCKGPLETDCCHQCAAGCTGPHNSDCLACLF 237  
 DB 195 PNGSGGGGEGENCOKLRIKCAOCCSHRCGRSPSDCHNCAAGCTGPRESDCLVQCKF 254  
 QY 238 NHSGICELHPALVTYNTDFFESMPNPEGRTFGASCTVCPYNYSTDVDS 289  
 DB 255 QDARTCKDTCPLMLNPTTYQMDVNPBGKYSFGATCVKCPKRYVYTDHGS 306

RESULT 7  
 ID ERBA HUMAN STANDARD; PRT: 1308 AA.  
 AC Q15303;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)  
 GN (p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).  
 OS ERBB4 OR HER4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM JM-A).  
 RC TISSUE-Breast carcinoma;  
 RX MEDLINE=93189574; PubMed=8383326;  
 RA Plozman G.D., Culouscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,  
 Foy L., Neubauer M.G., Shoyab M.;

RT "Ligand-specific activation of HER4/p180erbB4, a fourth member of the  
 RT epidermal growth factor receptor family.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).  
 RL [2]  
 RN SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).  
 RP TISSUE-Fetal brain;  
 RC MEDLINE=97476287; PubMed=9334263;  
 RX Elenius K., Corfas G., Choi C.J., Rio C., Plozman G.D.,  
 RA Klagsbrun M.;  
 RT "A novel transmembrane domain isoform of HER4/erbB4. Isoform-specific  
 RT tissue distribution and differential processing in response to  
 RT phorbol ester.";  
 RL J. Biol. Chem. 272:26761-26768(1997).  
 CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-  
 CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND  
 CC NTAF. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.  
 CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein.  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB  
 CC RECEPTORS (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: JM-A (SHOWN HERE) AND JM-B;  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER  
 CC FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED  
 CC BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND  
 CC NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN  
 CC CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,  
 CC KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,  
 CC PITUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,  
 CC LUNG, SALIVARY GLAND, AND PANCREAS.  
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).

CC EMBL: L07868; AAB59446.1; -;  
 CC HSSP: P11362; IEGK.  
 CC DR GeneW: HGNC:3432; ERBB4.  
 CC MIM: 600543; -;  
 CC DR InterPro: IPR000494; EGFR\_L\_domain.  
 CC DR InterPro: IPR000719; Euk\_pkinase.  
 CC DR InterPro: IPR002174; Furin-like.  
 CC DR InterPro: IPR001245; Tyr\_pkinase.  
 CC DR InterPro: IPR004019; YLP\_motif.  
 CC pfam: PF00069; pkinase.1.  
 CC pfam: PF00757; Furin-like.1.  
 CC pfam: PF01030; Recep\_L\_domain.2.  
 CC pfam: PF02757; YLP.2.  
 CC DR ProDom: PD000001; Euk\_pkinase.1.  
 CC DR SMART: SM00261; FU.4.  
 CC DR SMART: SM00219; TYRK.1.  
 CC DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 CC DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR.1.  
 CC DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 CC DR Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 CC Transmembrane; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 CC Alternative splicing.  
 CC KW SIGNAL.  
 CC FT CHAIN 1 25 POTENTIAL.  
 CC FT CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4.  
 CC FT DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).  
 CC FT TRANSMEM 652 675 POTENTIAL.  
 CC FT DOMAIN 676 1308 CYTOPLASMIC (POTENTIAL).  
 CC FT DOMAIN 186 334 CYS-RICH.  
 CC FT DOMAIN 496 633 CYS-RICH.



```

FT DOMAIN 718 985 PROTEIN KINASE.
FT NP_BIND 724 732 ATP (BY SIMILARITY).
FT BINDING 751 751 ATP (BY SIMILARITY).
FT ACT_SITE 843 843 BY SIMILARITY.
FT DISULFID 189 197 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 213 221 BY SIMILARITY.
FT DISULFID 217 229 BY SIMILARITY.
FT DISULFID 230 238 BY SIMILARITY.
FT DISULFID 234 246 BY SIMILARITY.
FT DISULFID 249 258 BY SIMILARITY.
FT DISULFID 262 289 BY SIMILARITY.
FT DISULFID 293 304 BY SIMILARITY.
FT DISULFID 308 323 BY SIMILARITY.
FT DISULFID 326 330 BY SIMILARITY.
FT DISULFID 503 512 BY SIMILARITY.
FT DISULFID 507 520 BY SIMILARITY.
FT DISULFID 523 532 BY SIMILARITY.
FT DISULFID 536 552 BY SIMILARITY.
FT DISULFID 555 569 BY SIMILARITY.
FT DISULFID 559 577 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 593 614 BY SIMILARITY.
FT DISULFID 617 625 BY SIMILARITY.
FT DISULFID 621 633 BY SIMILARITY.
FT MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 495 495 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 548 548 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 626 648 NGPISHDCTTPWGHSTLPQHA -> IGSTIEDCTGLMD
      (IN ISOFORM JM-B).
SQ SEQUENCE 1308 AA; 146807 MW; 5E4AE80985D88761 CRC64;

Query Match 45.1%; Score 716.5; DB 1; Length 1308;
Best Local Similarity 48.6%; Pred. NO. 5.9e-53;
Matches 138; Conservative 37; Mismatches 98; Indels 11; Gaps 3;

```

```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).
GN ERBB4 OR TYRO-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=98221155; PubMed=9553078;
RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,
RA Marchionni M.A., Kelly R.A.;
RT "Neuregulins promote survival and growth of cardiac myocytes.
RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult
RT ventricular myocytes."
RL J. Biol. Chem. 273:10261-10269(1998).
RN [2]
RP SEQUENCE OF 848-901 FROM N.A.
RC TISSUE=Sciatic nerve;
RX MEDLINE=91222560; PubMed=2025425;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system."
RL Neuron 6:691-704(1991).
RN [3]
RP SEQUENCE OF 1031-1198 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spinal cord;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Froehner P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuregulins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration."
RL J. Neurosci. 17:1642-1659(1997).
RN [4]
RP FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
RN NTRK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION,
RN NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB
CC RECEPTORS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING
CC NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS
CC OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE
CC RETICULAR NUCLEUS OF THE THALAMUS. VERY LOW LEVELS IN KIDNEY, AND
CC HEART.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF041838; AAD08899.1; -
DR EMBL: U52531; AAC53051.1; -
DR HSP: P11362.1; ICGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF02757; YLP; 2.

```

```

DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00219; TYKIC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL; 1 25
FT CHAIN; 26 1308
FT DOMAIN; 26 651
FT TRANSMEM; 652 675
FT DOMAIN; 676 1308
FT DOMAIN; 186 334
FT DOMAIN; 496 633
FT DOMAIN; 718 985
FT NP_BIND; 724 732
FT BINDING; 751 751
FT ACT_SITE; 843 843
FT DISULFID; 189 197
FT DISULFID; 193 205
FT DISULFID; 213 221
FT DISULFID; 217 229
FT DISULFID; 230 238
FT DISULFID; 234 246
FT DISULFID; 249 258
FT DISULFID; 262 289
FT DISULFID; 293 304
FT DISULFID; 308 323
FT DISULFID; 326 330
FT DISULFID; 303 312
FT DISULFID; 507 520
FT DISULFID; 536 552
FT DISULFID; 555 569
FT DISULFID; 577 577
FT DISULFID; 580 589
FT DISULFID; 617 625
FT DISULFID; 621 633
FT MOD_RES; 1162 1162
FT MOD_RES; 1188 1188
FT MOD_RES; 1258 1258
FT MOD_RES; 1284 1284
FT CARBOHYD; 138 138
FT CARBOHYD; 174 174
FT CARBOHYD; 253 253
FT CARBOHYD; 358 358
FT CARBOHYD; 410 410
FT CARBOHYD; 473 473
FT CARBOHYD; 495 495
FT CARBOHYD; 548 548
FT CARBOHYD; 576 576
FT CARBOHYD; 620 620
FT CONFLICT; 1062 1062
FT CONFLICT; 1080 1082
SQ SEQUENCE
1308 AA; 146957 MW; D944B0996A08B41 CRC64;

Query Match      44.5%; Score 706.5; DB 1; Length 1308;
Best Local Similarity 48.2%; Pred. No. 4, 2e-52;
Matches 137; Conservative 37; Mismatches 99; Indels 11; Gaps 3;

```

```

DB 136 LKMLTEINGVYVONKFLCYADTTHMODIVRNPMPSNMTLVSTIGSGCGRCHNSCTG 195
OY 181 SRCWGESSEDCQSLRTVCAGCG-ARCKGGLPTDCHEQACAGCTGPKHSDCLALHFNH 239
DB 196 -RCMGEPENHCQTLRTVYCAEDCDGRYGPYVDCCHRCBAGCGSGKDTLFCACNNFND 254
OY 240 SGICEHLCPALVYNTDFESMPNPEGRTFGASCYACPYNL 283
DB 255 SGACVYQCQPTFVYNPFTFQLEHNFNKAYYGAFCVKKCPHNFV 298.

RESULT 9
ERB3_HUMAN
ID ERB3_HUMAN STANDARD; PRT; 1342 AA.
AC P21860;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)
DE (c-erbB3) (Tyrosine kinase-type cell surface receptor HER3).
GN ERB3 OR HER3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90083234; PubMed=2687875;
RA Kraus M.H., Issing W., Miki T., Popescu N.C., Aaronson S.A.;
RT "Isolation and characterization of ERB3, a third member of the
RT ERB3/epidermal growth factor receptor family: evidence for
RT overexpression in a subset of human mammary tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9193-9197(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90311312; PubMed=2164210;
RA Plozman G.D., Whitney G.S., Neubauer M.G., Green J.M., McDonald V.L.,
RA Todaro G.J., Shoyab M.;
RT "Molecular cloning and expression of an additional epidermal growth
RT factor receptor-related gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:4905-4909(1990).
RN [3]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE=Placenta;
RX MEDLINE=93282822; PubMed=7685162;
RA Katoh M., Yazaki Y., Sugimura T., Terao M.;
RT "c-erbB3 gene encodes secreted as well as transmembrane receptor
RT tyrosine kinase.";
RL Biochem. Biophys. Res. Commun. 192:1189-1197(1993).
CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEURGUINS AND NTRK.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG FORM) AND
CC SECRETED (SHORT FORM).
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS, A LONG TRANSMEMBRANE FORM AND A
CC SHORT SECRETED FORM, EXIST DUE TO ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EPITHELIAL TISSUES AND BRAIN.
CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).
CC -1- PHOSPHORYLATION: OVEREXPRESSED IN A SUBSET OF HUMAN MAMMARY TUMORS.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

CC or send an email to license@slsb-sib.ch).

CC -----

DR EMBL: M29366; AAA35790.1; -

DR EMBL: M34309; AAA35979.1; -

DR EMBL: S61953; AAB26935.1; -

DR PIR: A36223; A36223.

DR HSSP: P13362; 1FCG.

DR Genew: HGNC:3431; ERB3.

DR MIM: 190151; -

DR InterPro: IPR000494; EGFR\_L\_domain.

DR InterPro: IPR000719; Euk\_Pkinase.

DR InterPro: IPR001245; Furin-like.

DR Pfam: PF00069; Pkinase; 1.

DR Pfam: PF00757; Furin-like; 1.

DR Pfam: PF01030; Recep\_L\_domain; 2.

DR ProDom: PD000001; Euk\_Pkinase; 1.

DR SMART: SM00261; FU; 3.

DR SMART: SM00219; TYRC; 1.

DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.

DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; FALSE\_NEG.

DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.

DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;

KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;

KW Alternative splicing.

FT SIGNAL 1 19 POTENTIAL.

FT CHAIN 20 1342 RECEPTOR PROTEIN-TYROSINE KINASE ERB3-3.

FT DOMAIN 20 643 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 644 664 POTENTIAL.

FT DOMAIN 665 1342 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 709 966 PROTEIN KINASE.

FT NP\_BIND 715 723 ATP (BY SIMILARITY).

FT BINDING 742 742 ATP (BY SIMILARITY).

FT ACT\_SITE 834 834 BY SIMILARITY.

FT DISULFID 186 194 BY SIMILARITY.

FT DISULFID 190 202 BY SIMILARITY.

FT DISULFID 210 218 BY SIMILARITY.

FT DISULFID 214 226 BY SIMILARITY.

FT DISULFID 227 235 BY SIMILARITY.

FT DISULFID 231 243 BY SIMILARITY.

FT DISULFID 246 255 BY SIMILARITY.

FT DISULFID 259 286 BY SIMILARITY.

FT DISULFID 290 301 BY SIMILARITY.

FT DISULFID 305 320 BY SIMILARITY.

FT DISULFID 323 327 BY SIMILARITY.

FT DISULFID 500 509 BY SIMILARITY.

FT DISULFID 504 517 BY SIMILARITY.

FT DISULFID 520 529 BY SIMILARITY.

FT DISULFID 533 549 BY SIMILARITY.

FT DISULFID 552 565 BY SIMILARITY.

FT DISULFID 556 573 BY SIMILARITY.

FT DISULFID 576 585 BY SIMILARITY.

FT DISULFID 589 610 BY SIMILARITY.

FT DISULFID 613 621 BY SIMILARITY.

FT DISULFID 617 629 BY SIMILARITY.

FT CARBOHYD 126 126 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 250 250 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 353 353 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 408 408 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 414 414 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 437 437 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 469 469 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 522 522 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 556 556 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT CARBOHYD 616 616 N-LINKED (GLCNAC; . . .) (POTENTIAL).

FT VANSPLIC 141 183 EILSGVIEKNDKLCHEMDTDMRDIVARDIAEIVKNGR

FT VANSPLIC 141 183 SC -> GQFPMVPSGLPQAPQADYLDLDDRLTLSSSK

FT VANSPLIC 184 1342 VPVLAIV (IN SHORT ISOFORM).

FT CONFLICT 560 560 MISSING (IN SHORT ISOFORM).

FT CONFLICT 1064 1064 E -> G (IN REF. 2).

FT CONFLICT 1064 1064 E -> G (IN REF. 2).

FT SEQUENCE 1342 AA; 148097 MW; 7201E7F66CA374BD CRC64;

Query Match 43.2%; Score 686; DB 1; Length 1342;

Best Local Similarity 46.1%; Pred. No. 2,46-50;

Matches 131; Conservative 44; Mismatches 95; Indels 14; Gaps 5;

QY 1 STQVCTGDMKRLPASEPETHLMDIRLYXGCGVYQVQLELYTPNLSLFDIOEQ 60

DB 25 SQAVCPGTLCNGSLVYGDENQYQTLKLYERCEVVMGLIETLGNMDSLQWIRVET 84

QY 61 GYVLAHNOVQVPLQRIARVGTOLFEDNYALVLDNGDPLNNTPTGASPGGLRELO 120

DB 85 GYVLAHNEFTPLPLNLRVVGTOYDCKFAIFV-----LNTVT-----NSSHALRQR 135

QY 121 LRSLEILKGYLVLIQRPOLCYODTILMKDIFKKNQALFLIDNRSRACPSPMCKG 180

DB 136 LQGLLEILSGGYIEKKNKLCHEMDTDMRDIVARDIAEIVKNGR 192

QY 181 SRCWGESSEDCQSLTRIVCAGGC-ARCKGRLPTDCHEQCAAGCTGPHNSDCLAFHNN 239

DB 193 -RCWGPGESEDCQTLTKITCAFCQNGHCSPNPNQCHDECAGCGSPDTCFACRHRND 251

QY 240 SGICELHCPALVTYNTDFEESMPNPEGRVTEGASCVTACPYNYL 283

DB 252 SGACVPRCPPLVYNTKLTFLQLEPNHTYXYQYGVCAVASCPIHNV 295

RESULT 10

ERB3\_RAT

ID ERB3\_RAT STANDARD: PRT: 1339 AA.

AC Q62799; Q62955;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Receptor protein-tyrosine kinase erb3-3 precursor (EC 2.7.1.112)

DE (c-erbB3).

GN ERB3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_Taxid=10116;

RN [1]

RC STRAIN=Sprague-Dawley; TISSUE=Liver;

RX MEDLINE=96096535; PubMed=8522190;

RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;

RT "Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein.";

RL Gene 165:279-284(1995).

RN [2]

RP REVISIONS TO 85: 513 AND 565.

RA Hellyer N.J., Koland J.G.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 922-1097 FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Sciatic nerve;

RX MEDLINE=97184212; PubMed=9030624;

RA Carroll S.L., Miller M.L., Frohneert P.W., Kim S.S., Corbett J.A.;

RT "Expression of neurogins and their putative receptors, ErbB2 and ErbB3, is induced during Wallerian degeneration.";

RL J. Neurosci. 17:1642-1659(1997).

CC -1 FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAN.

CC -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein tyrosine phosphate.

CC -1 SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS (POTENTIAL).

CC -1 SUBCELLULAR LOCATION: Type I membrane protein.

CC -1 DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.

CC -1 PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF PHOSPHATIDYLINOSITOL 3-KINASE.

CC -1 SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: 029339; AAC28498.2; -  
 EMBL: 052530; AAC3050.1; -  
 HSSP: P11362; 1FGK.  
 InterPro: IPR000494; EGFR\_L.domain.  
 InterPro: IPR000719; Euk\_pkinase.  
 InterPro: IPR002174; Furin-like.  
 InterPro: IPR001245; Tyr\_pkinase.  
 Pfam: PF00069; pkinase; 1.  
 Pfam: PF00757; Furin-like; 1.  
 Pfam: PF01030; Recep\_L.domain; 2.  
 PRINTS: PR00109; TYRKINASE.  
 ProDom: PD000001; Euk\_pkinase; 1.  
 SMART: SM00261; FU; 5.  
 SMART: SM00219; TYRKC; 1.  
 PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 PROSITE: PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
 SIGNAL 1 19  
 CHAIN 20 1339  
 DOMAIN 20 643  
 TRANSMEM 644 662  
 DOMAIN 663 1339  
 DOMAIN 183 259  
 DOMAIN 707 964  
 NP\_BIND 713 721  
 BINDING 740 740  
 ACT\_SITE 832 832  
 DISULFID 186 194  
 DISULFID 190 202  
 DISULFID 210 218  
 DISULFID 214 226  
 DISULFID 227 235  
 DISULFID 231 243  
 DISULFID 246 255  
 DISULFID 259 266  
 DISULFID 290 301  
 DISULFID 305 320  
 DISULFID 323 327  
 DISULFID 500 509  
 DISULFID 504 517  
 DISULFID 520 529  
 DISULFID 533 549  
 DISULFID 556 573  
 DISULFID 576 585  
 DISULFID 589 610  
 DISULFID 613 621  
 DISULFID 617 629  
 DISULFID 126 136  
 CARBOHYD 250 250  
 CARBOHYD 353 353  
 CARBOHYD 408 408  
 CARBOHYD 414 414  
 CARBOHYD 437 437  
 CARBOHYD 469 469  
 CARBOHYD 522 522  
 CARBOHYD 566 566  
 CARBOHYD 616 616  
 CONFLICT 1028 1028  
 SEQUENCE 1339 AA; 147545 MW; 0AA5F2402BBDFLE CRC64;

Query Match 41.5%; Score 658; DB 1; Length 1339;  
 Best Local Similarity 44.7%; Pred. No. 5, 6e-48;  
 Matches 127; Conservative 42; Mismatches 101; Indels 14; Gaps 5;

QY 1 STGYCTGTDMKRLPSPSEPHLMDLRLHYGCOVGNLEFLTFLPYNASLSTFLDIOEVO 60  
 DB 25 SQAVCEPTLNGLSGTGDADQYOTLYKYECEVMNLEIVLTGHNADLSFLDMIREVT 84  
 QY 61 GYVLIANQVROYPLQRLRYRGTOULFEDVYALVALDNGPLNNTPTVYASPGGLRELO 120  
 DB 85 GYVLIANNEFSLVPLPLRYAVRGTOYDGFALFVM-----LNNNT---NSSHALQOLK 135  
 QY 121 LRSITELKSGVLIQRPOLCYOTDTLMKDIFFHKNQALTLTIDTNSRACHPCSPMCKG 180  
 DB 136 FTQLTETLSGVYIEKNDKLCMDTIDMDIVNR---GAEIYVKNNGANCPCHEVCCKG 192  
 QY 181 SRCWGESSEDCQSITRTVACAGC-ARCKGPLPDCCHGCAAGCTGPKSHSLACLHFNH 239  
 DB 193 -RCWGPDPDCCQLITKTKICAPQCNRCFGFPNQCCHDEACAGCSGPDPTDCEACRRFND 251  
 QY 240 SGICELHCPALVYNTDFESMRPBERYFGASCVTACRYNLT 283  
 DB 252 SGACVPRCPPELVYNNKLTFLQLEPNPHTKYOGGVACVACPHNFV 295

RESULT 11  
 XMRK\_XIPMA  
 ID XMRK\_XIPMA STANDARD; PRT; 1167 AA.  
 AC P13388;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Melanoma receptor protein-tyrosine kinase precursor (EC 2.7.1.112).  
 GN XMRK OR TV.  
 OS Xiphophorus maculatus (Southern platyfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;  
 OC Cyprinodontiformes; Poeciliidae; Xiphophorus.  
 NX NCBI\_Taxid=8083;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90015140; PubMed=2797166;  
 RA Witbrodt J., Adam D., Malitschek B., Mauelet W., Raulf F.,  
 RA Telling A., Robertson S.M., Scharlt M.;  
 RT "Novel putative receptor tyrosine kinase encoded by the melanoma-  
 inducing Tu locus in Xiphophorus.";  
 RL Nature 341:415-421(1989).  
 RN [2]  
 RP REVISION TO 515.  
 RA Scharlt M.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DISEASE: INVOLVED IN PIGMENT CELLS MALIGNANT MELANOMAS.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X16891; CAA34770.2; -  
 PIR: S06142; S06142.  
 HSSP: P11362; 1FGK.  
 InterPro: IPR000494; EGFR\_L.domain.  
 InterPro: IPR000719; Euk\_pkinase.  
 InterPro: IPR002174; Furin-like.  
 InterPro: IPR002290; Ser\_thr\_pkinase.  
 InterPro: IPR001245; Tyr\_pkinase.  
 Pfam: PF00069; pkinase; 1.  
 Pfam: PF00757; Furin-like; 1.

```

DR Pfam: PF01030; Recep_L-domain: 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_Pkinase: 1.
DR SMART: SM00261; FU: 5.
DR SMART: SM00220; S_Tkc: 1.
DR SMART: SM00219; TyKc: 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR: 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM: 1.
KM Transmembrane; Glycoprotein; Receptor; Signal; Transferase.
FT Tyrosine-protein kinase; ATP-binding; Phosphorylation; Proto-oncogene.
FT SIGNAL 1 25
FT CHAIN 26 1167
FT DOMAIN 26 642 MELANOMA RECEPTOR PROTEIN-TYROSINE
FT TRANSMEM 643 665 KINASE.
FT DOMAIN 666 1167 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 710 977 POTENTIAL.
FT NE_BIND 716 724 CYTOPLASMIC (POTENTIAL).
FT BINDING 743 743 ATP (BY SIMILARITY).
FT ACT_SITE 835 835 ATP (BY SIMILARITY).
FT DISULFID 199 204 BY SIMILARITY.
FT DISULFID 220 212 BY SIMILARITY.
FT DISULFID 224 226 BY SIMILARITY.
FT DISULFID 237 245 BY SIMILARITY.
FT DISULFID 241 253 BY SIMILARITY.
FT DISULFID 256 265 BY SIMILARITY.
FT DISULFID 269 296 BY SIMILARITY.
FT DISULFID 300 311 BY SIMILARITY.
FT DISULFID 315 330 BY SIMILARITY.
FT DISULFID 333 337 BY SIMILARITY.
FT DISULFID 504 513 BY SIMILARITY.
FT DISULFID 508 521 BY SIMILARITY.
FT DISULFID 537 553 BY SIMILARITY.
FT DISULFID 556 569 BY SIMILARITY.
FT DISULFID 560 577 BY SIMILARITY.
FT DISULFID 593 615 BY SIMILARITY.
FT DISULFID 618 626 BY SIMILARITY.
FT DISULFID 622 634 BY SIMILARITY.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA; 129934 MW; 4793E0749DC1D55A CRC64;

Query Match 40.4%; Score 640.5; DB 1; Length 1167;
Best Local Similarity 44.2%; Pred. No. 1.4e-66;
Matches 126; Conservative 40; Mismatches 112; Indels 7; Gaps 4;

```

```

QY 242 ICHELCPALVYNTDFESMPNPEGRYTFGASCVTACPYNYLSTD 286
Db 264 TCRCTCPKRYDILVSHQVNDNPNIKYTFGACVACKECPSNVYVTE 308

RESULT 12
ID EGRF_DROME STANDARD: PRT; 1426 AA.
AC P04412; O61601; Q9W2G0; P81868;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermal growth factor receptor precursor (EG 2.7.1.112) (Egfr)
DE (Gurken receptor) (Torpedo protein) (Drosophila relative of ERBB).
GN EGRF OR TOP OR C-ERBB OR DER OR CG10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).
RX MEDLINE=94350209; Pubmed=8070664;
RA Clifford R., Schupbach T.;
RT "Molecular analysis of the Drosophila EGF receptor homolog reveals
RT that several genetically defined classes of alleles cluster in
RT subdomains of the receptor protein.";
RL Genetics 137:531-550(1994).
RN [2]
RP REVISIONS.
RA Clifford R., Schupbach T.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=85124611; Pubmed=2982499;
RA Livneh E., Glazer L., Segal D., Schlesinger J., Shilo B.-Z.;
RT "The Drosophila EGF receptor gene homolog: conservation of both
RT hormone binding and kinase domains.";
RL Cell 40:599-607(1985).
RN [4]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RC STRAIN-Oregon-R; TISSUE=Embryo;
RX MEDLINE=87002474; Pubmed=3093080;
RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;
RT "Alternative 5' exons and tissue-specific expression of the
RT Drosophila EGF receptor homolog transcripts.";
RL Cell 46:1091-1101(1986).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM TYPE I), TISSUE SPECIFICITY, AND MUTATION
RP ANALYSIS.
RX MEDLINE=99102120; Pubmed=9882502;
RA Lesokhin A.M., Yu S.-Y., Katz J., Baker N.E.;
RT "Several levels of EGF receptor signaling during photoreceptor
RT specification in wild-type, Ellipse, and null mutant Drosophila.";
RL Dev. Biol. 205:129-144(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM TYPE I).
RC STRAIN-Berkeley;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Plamkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadeu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N.S., Gebhart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mlshina N.V., Modarity C., Morris J., Mostreli A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacible J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*.";  
 RA Science 287:2185-2195(2000).  
 RN [7]  
 RP SEQUENCE OF 959-1078 FROM N.A.  
 RC STRAIN-Daekwanayoon;  
 RX MEDLINE=85137938; PubMed=2983232;  
 RA Wadsworth S.C., Vincent W.S. III, Blalodeau-Wentworth D.;  
 RT "A *Drosophila* genomic sequence with homology to human epidermal  
 RT growth factor receptor";  
 RL Nature 314:178-180(1985).  
 RN [8]  
 RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION  
 RP ANALYSIS.  
 RX MEDLINE=92038942; PubMed=1936959;  
 RA Raz E., Schejter E.D., Shilo B.Z.;  
 RT "Intracellular complementation among DER/fib alleles: implications for  
 RT the mechanism of signal transduction by receptor-tyrosine kinases";  
 RL Genetics 159:191-201(1991).  
 RN [9]  
 RP REVIEW.  
 RX MEDLINE=97248481; PubMed=9094709;  
 RA Perlmutter N., Perkins L.A.;  
 RT "There must be 50 ways to rule the signal: the case of the *Drosophila*  
 RT EGF receptor";  
 RL Cell 89:13-16(1997).  
 CC -1- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,  
 CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-  
 CC MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.  
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES. AND FOR THE  
 CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL  
 CC POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE  
 CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOSEROSA  
 CC AND VENTRAL NEUROECODERMAL CELLS. GERM BAND RETRACTION, CELL FATE  
 CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF  
 CC CUTICLE.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: ISOFORMS TYPE I AND TYPE II ARE TYPE I  
 CC MEMBRANE PROTEINS. ISOFORM TYPE III IS A TYPE III MEMBRANE  
 CC PROTEIN.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS, TYPE I (SHOWN HERE), TYPE II AND  
 CC TYPE III; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED IN EMBRYOS. IN LARVAE,  
 CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGE OF  
 CC TESTIS AND OVARY. AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST  
 CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW. LEVELS REMAIN HIGH  
 CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN  
 CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX  
 CC AND THORACIC AND ABDOMINAL GANGLIA.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL: AF052754; AAC08536.1; -  
 DR EMBL: AF052753; AAC08536.1; JOINED.  
 DR EMBL: AF052754; AAC08535.1; -  
 DR EMBL: AF052752; AAC08535.1; JOINED.  
 DR EMBL: K03054; AAA51462.1; -  
 DR EMBL: K03417; AAA51460.1; -  
 DR EMBL: K03416; AAA50965.1; -  
 DR EMBL: K03418; AAA51461.1; -  
 DR EMBL: AF109077; AAD26134.1; -  
 DR EMBL: AF109078; AAD26132.1; -  
 DR EMBL: AF109082; AAD26132.1; JOINED.  
 DR EMBL: AF109078; AAD26133.1; -  
 DR EMBL: AF109084; AAD26133.1; JOINED.  
 DR EMBL: AF109079; AAD26130.1; -  
 DR EMBL: AF109081; AAD26130.1; JOINED.  
 DR EMBL: AF109079; AAD26131.1; -  
 DR EMBL: AF109083; AAD26131.1; JOINED.  
 DR EMBL: AF109080; AAD26135.1; -  
 DR EMBL: AE003454; AAF46732.1; -  
 DR EMBL: X02293; CAA26157.1; -  
 DR EMBL: X78920; CAA55523.1; -  
 DR EMBL: X78918; CAA55521.1; -  
 DR EMBL: X78919; CAA55522.1; -  
 DR PIR: A00640; GOFEE.  
 DR HSP: P11362; IFGK.  
 DR FLYBASE: FBgn0003731; Egrf.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; EV\_7.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Receptor; Phosphorylation; Transferase;  
 KW Tyrosine-protein kinase; ATP-binding; Signal; Alternative splicing;  
 KW Developmental protein.  
 FT SIGNAL 1 30  
 FT CHAIN 31 1426  
 FT DOMAIN 31 868  
 FT TRANSMEM 869 889  
 FT DOMAIN 890 1426  
 FT DOMAIN 938 1198  
 FT NP\_BIND 944 952  
 FT BINDING 971 971  
 FT ACT\_SITE 1063 1063  
 FT MOD\_RES 902 902  
 FT PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 Query Match 33.7%; Score 535; DB 1; Length 1426;  
 Best Local Similarity 37.3%; Pred. No. 1,6e-37;  
 Matches 110; Conservative 43; Mismatches 116; Indels 26; Gaps 6;  
 QY 3 QVCTGDKMLRLPASETHLDMRLHYCGGVQGNLELYLP-NASLSFLQDIQEVG 61  
 DB 100 KIDIGTSRSLVSPNSNEHHRNRLDRYCTGVGDKMLTLPNENLDSFLDNIRVYTG 159  
 QY 62 YVLIANOVQVQLORLRIYRGOLF-----EDNVALLVNDGDPINNTPTVPGASGGL 116  
 DB 160 YLISHVDAKVKVPRQIIRGRTLSLSVEEKIALV-----TYSKM 203



```

FT DISULFID 791 805 POTENTIAL.
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 368 368 C -> Y (IN SY10).
FT VARIANT 469 469 C -> R (IN MN216).
FT VARIANT 700 700 C -> W (IN MN23).
FT VARIANT 753 753 C -> Y (IN SY11).
FT VARIANT 1065 1065 T -> I (IN SY16).
FT VARIANT 1074 1074 G -> E (IN SY7).
SQ SEQUENCE 1323 AA; 150510 MW; 6B0307EE53EEFA99 CRC64;

Query Match 25.0%; Score 396.5; DB 1; Length 1323;
Best Local Similarity 28.8%; Pred. No. 86-26;
Matches 89; Conservative 54; Mismatches 113; Indels 53; Gaps 9;

OY 4 VCSTGDMKRLPASPETHLMDLRLHYOGCVVGNLELYLPN----- 47
DB 39 LCGSTNGISRYGTGNI-LEDELTWRCGRVYGNLEITMENETKKMRETSNSTYDPK 97
OY 48 -----ASLSFLDIOEVGVLIANNOVROLRLRYGTQLEEDNALAVLDNGDP 101
DB 98 NEDSPKSIINFNDLEIRSLIYRANIOKISFPRIRVYGVDEYVHDN-ALYIHNKDK- 155
OY 102 LNNTPVTGASPGLERELOLRSLTEILKGVLIQRNPOLCY-ODTLIMKDFIKNNOLAL 160
DB 156 -----VHEVYMRRLRVIRNGSVTTQDNPKKCYIDKIDMKLTYDPD--VO 199
OY 161 TLIDTNRSRACH-----PCSPCKGSGRCWGSSEDCOSLRTVYACAGGCARC--KGPL 210
DB 200 KYETJSHQHCYONGKSMACHESC-NDKCSGSGDNDCCGVYRSVCPKSCQCFYSNSTS 258
OY 211 PTDCCHGCAAGCTGPKRHSCLACLFHNSGICELHCPALVYNTDTFESMPNREGRTYR 270
DB 259 SYECCDSACLGCTGHPKNCIACSKYEIDGICETCPSKRIRHGTGLRVFNDGRYON 318
OY 271 GASCVTACP 279
DB 319 GNHCVEKCP 327

RESULT 14
ILPR_BRALA
ID ILPR_BRALA STANDARD; PRT: 1363 AA.
AC 002466;
DB 01-NOV-1997 (Rel. 35, Created)
DB 01-NOV-1997 (Rel. 35, Last sequence update)
DB 15-JUN-2002 (Rel. 41, Last annotation update)
DE Insulin-like peptide receptor precursor (EC 2.7.1.112) (ILP receptor).
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OC NCBI_Taxid=7740;
OX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=66408719; PubMed=8813726;
RA Pashmiforush M., Chan S.J., Steiner D.F.:
RA "Structure and expression of the insulin-like peptide receptor from
RT amphioxus.";
RT Mol. Endocrinol. 10:857-866(1996).
CC -1- FUNCTION: THIS RECEPTOR BINDS TO THE INSULIN RELATED PEPTIDE AND
CC -1- HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: PROBABLE TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY
CC DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF
CC THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE
CC DOMAIN (BY SIMILARITY).

```

```

CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
CC RECEPTOR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S83394; AAB50848.1; -.
DR HSSP: P06213; IIRK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR003962; FcR1L_repeat.
DR InterPro: IPR002174; FcR1L_lik.
DR InterPro: IPR002011; RTK_kinaseII.
DR InterPro: IPR01245; Tyr_pkinase.
DR Pfam: PF00041; fn3; 3.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00757; FcR1L_lik; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00014; FcR1L_lik.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00060; FN3; 3.
DR SMART: SM00261; FU; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; ATP-binding; Phosphorylation; Signal.
FT SIGNAL 1 29
FT CHAIN 30 716
FT PROPEP 717 720
FT CHAIN 721 1363
FT DOMAIN 721 928
FT TRANSMEM 929 949
FT DOMAIN 950 1363
FT NP_BIND 994 1283
FT BINDING 1000 1008
FT BINDING 1028 1028
FT ACT_SITE 1148 1148
FT MOD_RES 1174 1174
FT CARBOHYD 51 51
FT CARBOHYD 97 97
FT CARBOHYD 137 137
FT CARBOHYD 278 278
FT CARBOHYD 483 483
FT CARBOHYD 599 599
FT CARBOHYD 617 617
FT CARBOHYD 665 665
FT CARBOHYD 666 666
FT CARBOHYD 711 711
FT CARBOHYD 732 732
FT CARBOHYD 736 736
FT CARBOHYD 743 743
FT CARBOHYD 816 816
FT CARBOHYD 885 885
FT CARBOHYD 898 898
SQ SEQUENCE 1363 AA; 154104 MW; 238120B4EAB1ED55 CRC64;

Query Match 19.9%; Score 315.5; DB 1; Length 1363;
Best Local Similarity 28.8%; Pred. No. 6-1e-19;
Matches 92; Conservative 39; Mismatches 109; Indels 79; Gaps 16;

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:49 ; Search time 19.3311 Seconds  
(Without alignments)  
3080.404 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STGVCTGTDMLRLPASPETH.....FGASCYACPNYLSTDVGS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP archaea:  
2: SP bacteria:  
3: SP fungi:  
4: SP human:  
5: SP invertebrate:  
6: SP mammal:  
7: SP mhc:  
8: SP organelle:  
9: SP phage:  
10: SP plant:  
11: SP rodent:  
12: SP virus:  
13: SP vertebrate:  
14: SP unclassified:  
15: SP virus:  
16: SP bacteriophage:  
17: SP archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	419	4 O9UK79	O9UK79 homo sapien
2	1450	91.4	1259	6 O18735	O18735 canis famill
3	739	46.6	527	13 O90836	O90836 gallus gall
4	723	45.6	149	6 O9B66	O9B66 oryctolagus
5	716.5	45.1	643	11 O9ERV6	O9ERV6 mus musculu
6	716.5	45.1	655	11 O9WVF5	O9WVF5 mus musculu
7	716.5	45.1	1210	11 O9EP98	O9EP98 mus musculu
8	708.5	44.6	1209	11 O9QX70	O9QX70 ratu
9	700.5	44.1	478	11 O9E5E0	O9E5E0 ratu
10	686	43.2	331	4 O9BUD7	O9BUD7 homo sapien
11	649.5	40.9	1165	13 O9YH40	O9YH40 xiphophorus
12	635.5	40.0	599	13 O9PSH2	O9PSH2 gallus gall
13	596.5	37.6	1328	13 P79754	P79754 fugu rubrip
14	543	34.2	1433	5 O9BIH9	O9BIH9 anopheles g
15	407	25.6	1137	13 O9W6F6	O9W6F6 gallus gall
16	398	25.1	150	6 O9B664	O9B664 oryctolagus

17	395	24.9	151	6 O9B665	O9B665 oryctolagus
18	377	23.8	1368	5 O23821	O23821 caenorhabdi
19	341	21.5	366	5 O26569	O26569 schistosoma
20	341	21.5	1717	5 O26566	O26566 schistosoma
21	331	20.9	334	5 O26567	O26567 schistosoma
22	331	20.9	342	5 O26568	O26568 schistosoma
23	317	20.0	1193	5 O9Y1X8	O9Y1X8 ephydaria f
24	289	18.2	1671	5 O9NUV5	O9NUV5 biomphalaria
25	287.5	18.1	1472	5 O9U5A8	O9U5A8 bombyx mori
26	273.5	17.2	89	11 O88459	O88459 mus musculu
27	271.5	17.1	1418	13 O93457	O93457 scophthalmu
28	271.5	17.1	1418	13 O8UW83	O8UW83 paralicthys
29	268	16.9	1749	5 O8T0W6	O8T0W6 echinococu
30	259.5	16.4	1412	13 O8UW84	O8UW84 paralicthys
31	256.5	16.2	1362	13 O9PV24	O9PV24 xenopus lae
32	252.5	16.1	2144	5 O9VD94	O9VD94 drosophila
33	252.5	15.9	1369	13 O8UW86	O8UW86 paralicthys
34	250	15.8	1358	13 O73798	O73798 xenopus lae
35	244	15.4	1368	13 O8UW85	O8UW85 paralicthys
36	232.5	14.7	1371	11 O9QVW4	O9QVW4 ratu
37	228.5	14.4	410	11 O63720	O63720 ratu
38	228.5	14.4	469	11 O63721	O63721 ratu
39	218.5	13.8	1245	13 O9TGH8	O9TGH8 scophthalmu
40	206.5	13.0	946	5 O9VJ04	O9VJ04 drosophila
41	205	12.9	868	5 O9VFE2	O9VFE2 drosophila
42	191	12.0	82	6 O9NOK4	O9NOK4 sus scrofa
43	189	11.9	131	5 O9BH16	O9BH16 anopheles g
44	185	11.7	1843	5 O968Y9	O968Y9 caenorhabdi
45	185	11.7	1846	5 O16131	O16131 caenorhabdi

## ALIGNMENTS

## RESULT 1

ID	Q9UK79	PRELIMINARY:	PRT:	419 AA.
AC	O9UK79;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	HERstatin.			
GN	HER-2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99415951; PubMed=10485918;			
RA	Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;			
RT	"The HER-2/neu receptor tyrosine kinase gene encodes a secreted			
RL	autoclnhibitor.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RA	Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Hemner W.D.;			
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, AF177761, AAD56009.2; -			
DR	InterPro; IPR000494; EGFR_L-domain.			
DR	InterPro; IPR002174; Furin-Like.			
DR	Pfam; PF00757; Furin-Like; 1.			
DR	Pfam; PF01030; Recep_L-domain; 1.			
DR	SMART; SM00261; FU; 1			
SO	SEQUENCE 419 AA; 45472 MW; FFC1BE347E2D030C CRC64;			

QY	1 STGVCTGTDMLRLPASPETHLMDLRLHYGCGVQVQGLLELYLPTNASISFLDIOEYQ 60
DB	22 STGVCTGTDMLRLPASPETHLMDLRLHYGCGVQVQGLLELYLPTNASISFLDIOEYQ 81

Oy	61	GVLIAHQNOVOPVQRIYRIGQOLFEDNVALVLNDGPNLNTPTVTGSPGSELEQ	120
Dd	82	GVLIAHQNOVOPVQRIYRIGQOLFEDNVALVLNDGPNLNTPTVTGSPGSELEQ	141
Oy	121	LRLSLTEILKGVLIORNPOLCYODTILMKDIFHHNNOLALTLIDTNSRACHPCSPCKG	180
Dd	142	LRLSLTEILKGVLIORNPOLCYODTILMKDIFHHNNOLALTLIDTNSRACHPCSPCKG	201
Oy	181	SRMCWESSEDDOQLTRYCAGGCAKCKGRLPTDCCHEGCAAGCTGPKHSDCLACLFHNHS	240
Dd	202	SRMCWESSEDDOQLTRYCAGGCAKCKGRLPTDCCHEGCAAGCTGPKHSDCLACLFHNHS	261
Oy	241	GICELHCALVLYNTDFESMPNNEGRTFCAACVTAACPYNVYSTDVG	289
Dd	262	GICELHCALVLYNTDFESMPNNEGRTFCAACVTAACPYNVYSTDVG	310

RESULT 2			
ID	018735	PRELIMINARY;	PRT: 1259 AA.
AC	018735;		
DT	01-JAN-1998 (TrEMBLrel. 05, Created)		
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Erbb-2.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_TaxId=9615;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yokota H.;		
RT	"cDNA cloning of erbb-2 from canine mammary gland";		
RL	Submitted (OCT-1987) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AB008451; BAA23127.1; ..		
DR	HSSP; P11362; IFGK.		
DR	InterPro; IPR002048; EF-hand.		
DR	InterPro; IPR000494; EGFR_L_domain.		
DR	InterPro; IPR00719; Euk.pkinase.		
DR	InterPro; IPR002174; Furlin-like.		
DR	InterPro; IPR001245; Tyr.pkinase.		
DR	InterPro; IPR004019; YLP_motif.		
DR	Pfam; PF00757; Furin-like; 1.		
DR	Pfam; PF00069; Recept_Like; 1.		
DR	Pfam; PF01030; Recept_L_domain; 2.		
DR	Pfam; PF02757; YLP; 2.		
DR	ProDom; PD000001; Euk_pkinase; 1.		
DR	SMART; SM00261; FUF; 3.		
DR	SMART; SM00219; TYRKc; 1.		
DR	PROSITE; PS00018; EF_HAND; UNKNOWN_1.		
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.		
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.		
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.		
DR	APP-binding; Transferase; Tyrosine-protein kinase.		
SO	SEQUENCE 1259 AA; 137989 MW; E37364d49c4acd46 CRC64;		

Query Match	91.48;	Score 1450;	DB 6;	Length 1259;
Best Local Similarity	91.38;	Pred. No. 1e-128;		
Matches 263;	Conservative	8;	Mismatches 17;	Indels 0;
			Gaps	0;

0Y	2	TQVCTGPTMKRLRASPETHLMDLRHLYOGGVVQGMLETLTYPTNASLFFLODIOEVG	61
Db	23	TQVCTGDMKRLRASPETHLMDLRHLYOGGVVQGMLETLTYPTNASLFFLODIOEVG	82
0Y	62	YVLIANNOVROVFORLRIVGTFQFEDNYALAVLDNGDPLNMTTPYTGASPGGLREIOL	121
Db	83	YVLIANSQVROIPQRLRLIVGTFQFEDNYALAVLDNGDPLLEGIPAPGAAGGGLREIOL	142
0Y	122	RSLEILKGGVLIQRNPOLCTQDTIILKKDIFHKNNQATLPLIDNRSAACHPCSPCKGS	181
Db	143	RSLEILKGGVLIQRNPOLCTQDTIILKKDVFHKNNQATLPLIDNRSAACHPCSPCKDA	202

QY 182 RCMGSSSEDCQSILRTFYAGCAGCARKGRLPDDCCHECCAAAGCTGPKRSDCLACIFHNSG 241

Db 203 HCMGSSSEDCQSILRTFYAGCAGCARKGRLPDDCCHECCAAAGCTGPKRSDCLACIFHNSG 262

QY 242 ICELHCPLALTYNNDTFESMNPBEGRTFFGASCTYACAPYNTLSDVS 289

Db 263 ICELHCPLALTYNNDTFESMNPBEGRTFFGASCTYACAPYNTLSDVS 310

	RESULT 3		
090836	ID	PRELIMINARY:	PRT; 527 AA.
090836	AC	Q90836.	
DT	01-NOV-1996	(TREMBREL_01, Created)	
DT	01-NOV-1996	(TREMBREL_01, Last sequence update)	
DT	01-JUN-2002	(TREMBREL_21, Last annotation update)	
DE	EGF/TGF-alpha receptor precursor.		
GN	C-ERBB.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
CC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RX	SEQUENCE FROM N.A.		
RX	MEDLINE=92123214; PubMed=1733751;		
RA	Flickinger T.W., Malhale N.J., Kung H.-J.;		
RT	"An alternatively processed mRNA from the avian c-erbB gene encodes a		
RT	soluble, truncated form of the receptor that can block ligand-		
RL	dependent transfection."		
RL	Mol. Cell. Biol. 12:883-893(1992).		
DR	EMBL; M77637; AAA46759.1; ..		
DR	InterPro; IPR000494; EGFR_LDomain.		
DR	InterPro; IPR002174; Furin-Like.		
DR	Pfam; PF00757; Furin-Like; 1.		
DR	Pfam; PF01030; Recep_LDomain; 2.		
DR	SMART; SMO0261; FU; 2.		
KW	Receptor; Signal.		
FT	SIGNAL		
FT	CHAIN	1 28	POTENTIAL.
SO	SEQUENCE	527 AA; 58353 MW; 764364ABCC095298 CRC64;	EGF/TGF-ALPHA RECEPTOR.

Query Match	46.68;	Score 739;	DB 13;	Length 527;
Best Local Similarity	48.18;	Pred. No. 1.1e-61;		
Matches 139; Conservative	48;	Mismatches 90;	Indels 12;	Gaps 5

Qy	3	QVCGTGDMLKRLPASETHNIDMLRHLXGCGOYMGCSNLETATLPRTNALSLEODIGEVOCY	62
Db	38	KVCGGTSNKKTLTGCHVEDHTFSLDQRNINCEVSLNSETIYVEHNRDLTFLTKTIOEAGI	97
Qy	63	VLLAHNOVROVPLQRLRIYRGTOLEFEDNYALVALDNGDPLNNTPTVYGASPGGLREIOLR	12
Db	98	VLLALNADVIDPLENQTIRGVNVLIDNSPALVLSNYH - NKKTG - - - - - GLREIPLMK	14
Qy	123	SLTFILKGVTLQIRNRQTLCTYOTLLIMKRODIFKKNQDLATLTD - TNRSRACHPSPMKCS	18
Db	149	RLLSTILNGVKVKSINNRKLCNMDPTVLMNDIIDSRR - PLTYLDPAFSLNUSPCCKHPCTED	20
Qy	182	RCWESSESDQSILTRTFRGAGGGA - ECKGRPLPTDCCHEGACAGTSPKXSDCLACLFPHNS	24
Db	208	HCWAGSQNGQTLTFLKVICAGQCSGRCKRGVAPDDCCHNQACAGCTGPRESDCLACRFRDD	26
Qy	241	GICELHCPALVYTNPTTFESMNPGRGRTTFGASCYTACPRYVLTSDVGS	289
Db	268	ATCKDTCRPLVLNPTTYQMDVNPBGKVSFGATQVRECPRHNVYVTDHGS	316

RESULT 4		
Q9BG66		
ID	Q9BG66	PRELIMINARY;
AC	Q9BG66:	PRT; 149 AA.
DT	01-JUN-2001 (TReMBLrel. 17, Created)	.
DT	01-JUN-2001 (TReMBLrel. 17, Last sequence update)	.

DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
 DE Receptor tyrosine kinase Erb2 (Fragment).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,  
 RA Tetens F., Fischer B.  
 RT "ErbB genes and epidermal growth factor- (EGF-) like ligands in the  
 RT peti-implantation rabbit uterus and blastocyst."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF333178; AAK14371.1;  
 DR InterPro: IPR002174; Furin-like.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR SMART: SM00261; FU; 2.  
 KM Kinase.  
 FT NON\_TER  
 FT TER  
 SQ SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;  
 Query Match 45.68; Score 723; DB 6; Length 149;  
 Best Local Similarity 81.28; Pred. No. 7.5e-61;  
 Matches 121; Conservative 8; Mismatches 20; Indels 0; Gaps 0;  
 QY 137 NPQCYDTILMKDIFHNQALTLIDTNRSAHPCSPKSGESSEDCQSLTR 196  
 DB 1 NPQCYDTILMKDIFHNQALTLIDTNRSAHPCSPKSGESSEDCQSLTR 60  
 QY 197 TVCAGGACRCKGRLPTDCHEOCAGCTGPRHSDCLAGHNSGICELHCPALVTYNTD 256  
 DB 61 TICAGGACRCKGRLPTDCHEOCAGCTGPRHSDCLAGHNSGICELHCPALVTYNTD 120  
 QY 257 TFESEMPNEGRYTCGASCTACPYNYLST 285  
 DB 121 TFESEMPNEGRYTCGASCTACPYNYLST 149  
 RESULT 5  
 QSERV6  
 ID QSERV6 PRELIMINARY; PRT; 643 AA.  
 AC QSERV6;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor isoform 2.  
 GN EGFR.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV7AC;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Mahle N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF275365; AAG28046.1; JOINED.  
 DR EMBL: AF275364; AAG28046.1; JOINED.  
 DR MGD: MGI:95294; Egfr.  
 DR InterPro: IPR000345; CytC\_heme\_bind.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR002174; Furin-like.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recept\_L\_domain; 2.  
 DR SMART: SM00261; FU; 4.

DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 KM Receptor.  
 SQ SEQUENCE 643 AA; 71476 MW; DEF22002C084911B1 CRC64;  
 Query Match 45.18; Score 716.5; DB 11; Length 643;  
 Best Local Similarity 47.38; Pred. No. 1.9e-59;  
 Matches 138; Conservative 39; Mismatches 96; Indels 19; Gaps 4;  
 QY 3 QVCGTDMKLLPASPENHLDMLRHLYGCGVQGVGNLETLTPNASTSFLQDIQVGY 62  
 DB 29 KVCQSTSRRLQGLTFEEDHFLSLQRMNCEVJLNLLETIVORYDLSFLKTIQVAGY 88  
 QY 63 VLIANVOYPLQRLRIVRGCTOLEFEDNYVALVDNGPLNNTTPYVTCASPGRLQLR 122  
 DB 89 VLIANVTERIPLENLQIRGNALYENTYALATLSN-----YGNRTGLRELPMR 138  
 QY 123 SLTELKGVLIQRPOLCYDTILMKDI---FHKNQALTLIDTNRSAHPCSPMC 178  
 DB 139 NLOETLIGAVRFSSNPILCNMDITQWRDIONVFNMSNMDL---OSHPSRCPCDPSC 194  
 QY 179 KGRQWGESSEDCQSLTRTVGAGGA--RCKGRLPTDCHEOCAGCTGPRHSDCLAGH 237  
 DB 195 PNGSCWGGEGENCQKRLKTIQAOCSHRGRSPSDCHNOCAGCTGPRHSDCLVCKF 254  
 QY 238 NMSGICELHCPALVTYNTDFESEMPNEGRYTCGASCTACPYNYLSTVGS 289  
 DB 255 QDEATCKRQTCPLMLYNTTYQMOVNPGRKYSFGATCYKCKPRMYVTVDHGS 306  
 RESULT 6  
 Q9WVF5  
 ID Q9WVF5 PRELIMINARY; PRT; 655 AA.  
 AC Q9WVF5;  
 DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor (Epidermal growth factor receptor  
 DE isoform 3).  
 GN EGFR.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LIVER.  
 RA Reiter J.L., Threadgill D.W., Danielson A.J., Scheil C.,  
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
 RA Mahle N.J.;  
 RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode  
 RT Carboxy-Terminal Truncated Receptors."  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEV7AC;  
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,  
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Mahle N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egfr transcripts encoding truncated receptor  
 RT isoforms."  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LIVER.  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,



DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor.  
 GN EGFR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RX MEDLINE=9025888; PubMed=2342466;  
 RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,  
 RA Earp H.S.;  
 RT "A truncated, secreted form of the epidermal growth factor receptor is  
 RT encoded by an alternatively spliced transcript in normal rat tissue."  
 RL Mol. Cell. Biol. 10:2973-2982(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RA Petch L.A.;  
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RA Guttridge K., Dawson T.L., Earp H.S.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; M37394; AAF14008.1; -.  
 DR HSSP; P11362; IEGK.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR Prodom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00261; Fv; 3.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DYR; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.  
 KW SEQUENCE 1209 AA; 134891 MW; 96FEF7F6C1B773 CRC64;

Query Match 44.68; Score 708.5; DB 11; Length 1209;  
 Best Local Similarity 47.48; Pred. No. 2.3e-58;  
 Matches 137; Conservative 41; Mismatches 98; Indels 13; Gaps 4;

QY 3 QVCTGDMKRLPASPETHLMDLRLHYGCGVOGNELETLPYNASISFLDIOEVGY 62  
 DB 29 KVCQGTNRRLTQLGTFEDHFLSLQRMNCEVVLGNLEITTYQRRYDLSFKLTIOEVGY 88  
 QY 63 VLIANOVQVPLQRLIRVGTOLFEEDNYALAVLDNGPDLNNTTPVYGASPGRLQLR 122  
 DB 89 FLIALNTERIPSEDLQIRGNALYENTYALALISN-----YGTNRTGLRELPMR 138  
 QY 123 SLTEILKGVLIQNRPOLCYODTILMKDIFHKNNQALTLIDTNRS-RACHPCSPMKGS 181  
 DB 139 NLQELLGAVRFSSNPPLICNNMTIOWRDI-VQVFLSMMSMDVORHLILGCPKCPSPCNG 197  
 QY 182 RCGWESSDQSLRTVYAGCA-RCKGRLPTDCHEOCAGCTGPKHSDCLACLHFNHS 240  
 DB 198 SCWGRGEGNCKLTKIICAOQCSRCRCRSPSDCHNCAGCTGPRSDCLVCHRFDE 257  
 QY 241 GICELHCPALVYNTDFESMPNPEGRTYFGASCVTACPYNYLSTVGS 289  
 DB 258 ATCQDTQPLMLYNPTTYQMDVNPCKGKYSFGATCVKCPRNYYVTVDHGS 306

RESULT 9

Q9ESE0  
 ID Q9ESE0 PRELIMINARY; PRT; 478 AA.  
 AC Q9ESE0;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor related protein.  
 GN ERPP.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=GASTRO-DUODENAL MUCOUS;  
 RA Yu Y., Koshier J.A., Majumdar A.P.N.;  
 RT "Cloning of a novel EGFR-related peptide: A putative negative  
 RT regulator of EGFR."  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF187818; AAG17037.2; -.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR002174; Furin-like.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 2.  
 KW Receptor.  
 KW SEQUENCE 478 AA; 53233 MW; CF873A8376C519E5 CRC64;

Query Match 44.18; Score 700.5; DB 11; Length 478;  
 Best Local Similarity 46.68; Pred. No. 4.3e-58;  
 Matches 136; Conservative 39; Mismatches 98; Indels 19; Gaps 4;

QY 3 QVCTGDMKRLPASPETHLMDLRLHYGCGVOGNELETLPYNASISFLDIOEVGY 62  
 DB 29 KVCQGTNRRLTQLGTFEDHFLSLQRMNCEVVLGNLEITTYQRRYDLSFKLTIOEVGY 88  
 QY 63 VLIANOVQVPLQRLIRVGTOLFEEDNYALAVLDNGPDLNNTTPVYGASPGRLQLR 122  
 DB 89 FLIALNTERIPSEDLQIRGNALYENTYALALISN-----YGTNRTGLRELPMR 138  
 QY 123 SLTEILKGVLIQNRPOLCYODTILMKDIFHKNNQALTLIDTNRSAACHPCSPMC 178  
 DB 139 NLQELLGAVRFSSNPPLICNNMTIOWRDIYQVPMNSMSML---GHPSSCPKCPSC 194  
 QY 179 KGSRCWESSDQSLRTVYAGCA-RCKGRLPTDCHEOCAGCTGPKHSDCLACLHFNHS 237  
 DB 195 PNGSCWGRGEGNCKLTKIICAOQCSRCRCRSPSDCHNCAGCTGPRSDCLVCHRFDE 254  
 QY 238 NMSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACPYNYLSTVGS 289  
 DB 255 QDEATCKDTQPLMLYNPTTYQMDVNPCKGKYSFGATCVKCPRNYYVTVDHGS 306

RESULT 10  
 Q9BUD7  
 ID Q9BUD7 PRELIMINARY; PRT; 331 AA.  
 AC Q9BUD7;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Similar to v-erb-b2 avian erythroblastic leukemia viral oncogene  
 DE homolog 3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC002706; AA02706.1; -.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR002174; Furin-like.

DR Pfam: PF00757; Furin-like. 1.  
DR Pfam: PF01030; Recep\_L\_domain. 1.  
DR SMART: SM00261; FU; 2.  
SQ SEQUENCE 331 AA; 36489 MW; 45B8EBE683FE7E8 CRC64;  
Query Match 43.2%; Score 686; DB 4; Length 331;  
Best Local Similarity 46.1%; Pred. No. 6.5e-57;  
Matches 131; Conservative 44; Mismatches 95; Indels 14; Gaps 5;  
QY 1 STQCTGDMKRLRPASPEHLDMLRHLVYOGCVVQGNLELYLPTNASLFLDIDQEVQ 60  
DB 25 SQAVCPETLNGSLVGTGAENQYQTLVLYERCEVYVKNLEIVLGHADSLFLQWIREVT 84  
QY 61 GYVLIANQVROVPLQRLRVKRGTOLEFEDNYALAVLDNGPLNTPYVPGASPGGLRELQ 120  
DB 85 GYVLIANNESTPLPLRLRVKRGVYDGFALFVW-----LVNNT-----NSSHALQLR 135  
QY 121 LRSITELKGVLIQNRPOCLCYODTILMKDIFHKNNQALALTLIDTNRSRACHPCSPMCKG 180  
DB 136 LTLQLETLSSGVYIEKDKLCHMDTIDMRIVDRD---AEIVYKDKNGSCPCHVEYCKG 192  
QY 181 SRCWGESSEDCQSLRTVACGGC-ARCKGRLPTDCHECCAGCTGPKHSDCLACLFHFNH 239  
DB 193 -RCWGPSEDCQTLTKTICAPQCGNGHCFGNPQCCHDECAGCGSPQDTDFACRHFND 251  
QY 240 SGICELHCPALVYNTDFESMNPREGRTFGASCYVACPYNL 283  
DB 252 SGACVPRCPQPLVYNNKLTFLQLEPMPHKYQGVGVACVACPHNFV 295  
RESULT 11  
Q9YH40 PRELIMINARY; PRT; 1165 AA.  
AC Q9YH40;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Receptor tyrosine kinase proto-oncogene.  
GN XKRR.  
OS Xiphophorus xiphidium.  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.  
OX NCBI\_TaxID=8086;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RIO PURIFICATION;  
RX MEDLINE-98241172; PubMed-9582016;  
RA Dimitrijevic N., Winkler C., Wellbrock C., Gomez A., Duschl J.,  
RA Altschmid J., Scharl M.;  
RT Activation of the xmrk proto-oncogene of Xiphophorus by  
RT overexpression and mutational alterations.";  
RL Oncogene 16:1681-1690(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RIO PURIFICATION;  
RA Scharl M.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U53471; AAD10500.2; -  
DR HSSP: P11362; 1EGK  
DR InterPro: IPR000345; CytC\_heme\_bind.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001899; Gram\_pos\_anchor.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00757; Furin-like. 1.  
DR Pfam: PF00069; pkinase. 1.  
DR Pfam: PF01030; Recep\_L\_domain. 2.  
DR PRINTS: PR00109; TYRKINASE.  
DR PRODOM: PD000001; Euk\_pkinase. 1.  
DR SMART: SM00261; FU; 3.

DR SMART: SM00219; Tyrc. 1.  
DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; UNKNOWN\_1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 2.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR ATP-binding; Kinase; transferase; tyrosine-protein kinase.  
KW SEQUENCE 1165 AA; 129614 MW; 7F7EE8DB771A74E CRC64;  
Query Match 40.9%; Score 649.5; DB 13; Length 1165;  
Best Local Similarity 44.6%; Pred. No. 8.7e-53;  
Matches 127; Conservative 40; Mismatches 111; Indels 7; Gaps 4;  
QY 3 QVCTGDMKRLRPASPEHLDMLRHLVYOGCVVQGNLELYLPTNASLFLDIDQEVQ 62  
DB 28 KVCQGTSMQMTM---LDNHLYLKKMKMYSGCNVYLENLEITYTEQNDLSFLQSIQEVGY 84  
QY 63 VLIANQVROVPLQRLRVKRGTOLEFEDNYALAVLDNGPLNTPYVPGASPGGLRELQ 122  
DB 85 VLIANNEVSTPLNRLNLKGNLYEGNFTLVNYSYQK-NPSP--DYVQVGLKQLLS 141  
QY 123 SLTEILKGVLIQNRPOCLCYODTILMKDIFHKNNQALALTLIDTNRSRACHPCSPMCKGS 182  
DB 142 NLTEILSSGVYVSHNPLLCVETINMWDYDKTSNPYMLILPHAFERQCKQDPGCVNGS 201  
QY 183 CWGESSEDCQSLRTVACGGC-ARCKGRLPTDCHECCAGCTGPKHSDCLACLFHFNHSG 241  
DB 202 CWAPGPGHCQKFTKLCAECCNRCRGPIDCCNEHCAGCTGPRATDCLACRFNDG 261  
QY 242 ICHELCPALVYNTDFESMNPREGRTFGASCYVACPYNLSD 286  
DB 262 TCKDTCPPKTIYDLSHQVVDNPNKITFGAACVKECPSNVYTE 306  
RESULT 12  
Q9PSH2 PRELIMINARY; PRT; 599 AA.  
AC Q9PSH2;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Epidermal growth factor receptor (Fragment).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-94020816; PubMed-8414496;  
RA Callaghan T., Antczak M., Flickinger T., Raines M., Myers M.,  
RA Kung H.J.;  
RT "A complete description of the EGF-receptor exon structure:  
RT implication in oncogenic activation and domain evolution.";  
RL Oncogene 8:2939-2948(1993).  
RN [2]  
RP InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR002174; Furin-like.  
DR Pfam: PF00757; Furin-like. 2.  
DR Pfam: PF01030; Recep\_L\_domain. 2.  
DR SMART: SM00261; FU; 3.  
SQ SEQUENCE 599 AA; 66363 MW; FEAB46D293D991BD CRC64;  
Query Match 40.0%; Score 635.5; DB 13; Length 599;  
Best Local Similarity 29.1%; Pred. No. 8.3e-52;  
Matches 139; Conservative 48; Mismatches 89; Indels 201; Gaps 6;  
QY 3 QVCTGDMKRLRPASPEHLDMLRHLVYOGCVVQGNLELYLPTNASLFLDIDQEVQ 62  
DB 5 KVCQGTSMQMTM---LDNHLYLKKMKMYSGCNVYLENLEITYTEQNDLSFLQSIQEVGY 84  
QY 63 VLIANQVROVPLQRLRVKRGTOLEFEDNYALAVLDNGPLNTPYVPGASPGGLRELQ 122  
DB 65 VLIANNEVSTPLNRLNLKGNLYEGNFTLVNYSYQK-NPSP--DYVQVGLKQLLS 141





QY 120 QLSRLEILGCVLIQNPOLCYQDTILMKDI-FHKNNQALALILIDNRSRACHPCSPMC 178  
DB 105 ELALRDLILGGSVGFENNVLCHMKSLINMEIILLAPOTSQMYTFNFSSPRVPPCPHSC 164  
QY 179 KGSRCMGESESDCOSLRTVCAGGCA--RCKGPLPTDCHEOCAGCTGPKHSDCLCLH 236  
DB 165 EVG-CMGEAHNCORFESKLCSPQCSGRCFCGRPRECHLFCAGAGCTGPTQSDCLACKN 223  
QY 237 FNHSGICELCPALVYNTDTFESMPNPEGRTYFGASCYTACPYNYLSTDVGS 289  
DB 224 FYDDGVCKQCECPMQIYNPTNYFMEHPNPDCKYVYGATCVRKCP-EHLKDNKA 275

## RESULT 15

Q9W6F6 PRELIMINARY; PRT; 1137 AA.  
ID Q9W6F6  
AC Q9W6F6;  
DT 01-NOV-1999 (T-EMBLrel. 12, Created)  
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE Receptor tyrosine kinase (Fragment).  
GN ERBB4.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HINDBRAIN;  
RX MEDLINE=99263203; PubMed=10328884;  
RA Dixon M., Lumsden A.;  
RT "Distribution of neuropilin-1 (np1) and erb4 transcripts in  
RL embryonic chick hindbrain.";  
EMBL: AF21963; AAD31764.1; -  
DR HSP; P11362; IFGK.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_kinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001368; TNFR-c6.  
DR InterPro: IPR001245; Tyr\_kinase.  
DR InterPro: IPR004019; YLP\_motif.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF02757; YLP; 2.  
DR Pfam: PF02757; Recep\_L\_domain; 1.  
DR PRINTS: PR00109; TYRKINASE.  
DR PRODOM: PD000001; Euk\_kinase; 1.  
DR SMART: SM00261; FU; 3.  
DR SMART: SM00219; TYRK; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
KW kinase, tyrosine-protein kinase.  
FT NON\_TER  
SQ SEQUENCE 1137 AA; 127927 MW; 4D616436F87DC84F CRC64;

Query Match 25.6%; Score 407; DB 13; Length 1137;  
Best Local Similarity 48.3%; Pred. No. 8.8e-30;  
Matches 70; Conservative 18; Mismatches 55; Indels 2; Gaps 2;

QY 140 LCYQDTILMKDIFHKNNQALALILIDNRSRACHPCSPMCKSRMGESSESDCOSLRTVC 199  
DB 3 LCFADTILHMODIYRNPWASNFLLVPTNGSSGCGRCHKSCGTG-RQWGPTEHNCOTLTKTVC 61  
QY 200 AGGC-ARCKGRLPTDCHEOCAGCTGPKHSDCLCLHFNHSGICELCPALVYNTDTF 258  
DB 62 AEOCDGRCYGPVYVSDCHRCAGCGSKPDTCFACMNFNDSGACVYQCPOTFVYNPPTF 121  
QY 259 ESMNPNEGRTYFGASCTACPYNYL 283

DB 122 QLEHNNNAKTYGAFCKKCPHNFV 146

Search completed: April 28, 2003, 13:41:35  
Job time: 23.3311 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 28, 2003, 13:39:54 ; Search time 8.45736 Seconds  
(without alignments)  
1005.423 Million cell updates/sec

Title: US-09-821-883-23  
Perfect score: 1587  
Sequence: 1 STGYCTGDMKRLRLPASPER.....FGASCYACPYNTLSTVGS 289

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCUTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfilest.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1587	100.0	419	US-09-630-155-2	Sequence 2, Appli
2	1587	100.0	624	US-08-422-108-1	Sequence 1, Appli
3	1587	100.0	624	US-08-422-734-1	Sequence 1, Appli
4	1587	100.0	782	US-09-146-283-4	Sequence 4, Appli
5	1587	100.0	782	US-08-579-823A-4	Sequence 4, Appli
6	1587	100.0	782	US-09-344-195-4	Sequence 4, Appli
7	1587	100.0	1235	US-08-467-083-68	Sequence 68, Appli
8	1587	100.0	1235	US-08-414-417B-68	Sequence 68, Appli
9	1587	100.0	1235	US-08-484-438-8	Sequence 68, Appli
10	1587	100.0	1235	US-08-486-348A-68	Sequence 68, Appli
11	1587	100.0	1235	US-08-625-101-2	Sequence 2, Appli
12	1587	100.0	1235	US-08-468-545B-68	Sequence 68, Appli
13	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
14	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
15	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
16	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
17	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
18	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
19	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
20	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
21	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
22	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
23	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
24	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
25	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
26	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli
27	1587	100.0	1235	US-08-336-786-2	Sequence 2, Appli

28	493	31.1	97	4	US-09-046-783-3	Sequence 3, Appli
29	244.5	15.4	1382	2	US-08-737-715-2	Sequence 2, Appli
30	244.5	15.4	1382	4	US-09-457-040B-7	Sequence 7, Appli
31	232.5	14.7	516	4	US-08-746-559A-4	Sequence 4, Appli
32	232.5	14.7	1367	2	US-08-249-687C-2	Sequence 2, Appli
33	232.5	14.7	1367	2	US-08-625-819-2	Sequence 2, Appli
34	232.5	14.7	1367	4	US-08-746-559A-2	Sequence 4, Appli
35	232.5	14.7	1367	4	US-08-864-641B-18	Sequence 18, Appli
36	231.5	14.6	486	3	US-08-746-559A-5	Sequence 5, Appli
37	186	11.7	366	4	US-08-857-076-103	Sequence 103, App
38	185	11.7	1724	4	US-08-857-076-12	Sequence 12, Appli
39	171.5	10.8	383	4	US-08-857-076-105	Sequence 105, App
40	166.5	10.5	370	4	US-08-857-076-104	Sequence 104, App
41	132	8.3	381	4	US-08-857-076-106	Sequence 106, App
42	108.5	6.8	2211	4	US-09-738-884-1	Sequence 1, Appli
43	103.5	6.5	769	2	US-08-789-078-1	Sequence 1, Appli
44	103.5	6.5	769	2	US-08-752-633-1	Sequence 1, Appli
45	103.5	6.5	769	2	US-08-476-062A-45	Sequence 45, Appli

## ALIGNMENTS

RESULT 1  
US-09-630-155-2  
Sequence 2, Application US/09630155  
Patent No. 6414130  
GENERAL INFORMATION:  
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton  
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DAVIS WRIGHT TREMAYNE LLP  
STREET: 1501 Fourth Avenue, 2600 Century Square  
CITY: Seattle  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: PC compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/630,155  
FILING DATE: 16-Jan-2001  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Davison, Barry L.  
REGISTRATION NUMBER: 47,309  
REFERENCE/DOCKET NUMBER: 49321-10  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206 628-7621  
TELEFAX: 206 628-7699  
SEQUENCE FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 419  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: polypeptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-630-155-2  
Query Match 100.0%; Score 1587; DB 4; Length 419;  
Best Local Similarity 100.0%; Pred. No. 3.6e-139;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 STGYCTGDMKRLRLPASPERTHLDMLRHLYOGGVQVGNLELYLPFNMSLFDIOEVO 60  
DB 22 STGYCTGDMKRLRLPASPERTHLDMLRHLYOGGVQVGNLELYLPFNMSLFDIOEVO 81  
QY 61 GYLLAHNOVROVPLQRLRVGTOLFEEDNVAVLAVLDNGDPLNNTTPTVGASPGGLRELO 120

Db 82 GYVLIANQVAPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQ 141  
Qy 121 LRSITELKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
Db 142 LRSITELKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201  
Qy 181 SRCGSESEDCQSLTRIVCAGGACRCGKPLPTDCHEQCAAGCTGPKHSCLACLFHNS 240  
Db 202 SRCGSESEDCQSLTRIVCAGGACRCGKPLPTDCHEQCAAGCTGPKHSCLACLFHNS 261  
Qy 241 GICELHCPALVTYNTDFESMPNEGRTYTGASCVTACPNYISTDVG 289  
Db 262 GICELHCPALVTYNTDFESMPNEGRTYTGASCVTACPNYISTDVG 310

RESULT 2  
US-08-422-108-1  
Sequence 1, Application US/08422108  
Patent No. 6015567  
GENERAL INFORMATION:  
APPLICANT: Hudziak, Robert M.  
APPLICANT: Shepard, H. Michael  
APPLICANT: Ullrich, Axel  
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,108  
FILING DATE: 14-Apr-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/355460  
FILING DATE: 13-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/048346  
FILING DATE: 15-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/354319  
FILING DATE: 19-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 554C2D2  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-422-108-1

Query Match 100.0%; Score 1587; DB 3; Length 624;  
Best Local Similarity 100.0%; Pred. No. 6e-139;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STGVCTGTDKRLRLPASPEFHLDMLRLHYGCGVVGAGNLELTYLPTNASLSFLQDIOEYQ 60  
Db 1 STGVCTGTDKRLRLPASPEFHLDMLRLHYGCGVVGAGNLELTYLPTNASLSFLQDIOEYQ 60

Qy 61 GYVLIANQVAPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120  
Db 61 GYVLIANQVAPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120  
Qy 121 LRSITELKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
Db 121 LRSITELKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
Qy 181 SRCGSESEDCQSLTRIVCAGGACRCGKPLPTDCHEQCAAGCTGPKHSCLACLFHNS 240  
Db 181 SRCGSESEDCQSLTRIVCAGGACRCGKPLPTDCHEQCAAGCTGPKHSCLACLFHNS 240  
Qy 241 GICELHCPALVTYNTDFESMPNEGRTYTGASCVTACPNYISTDVG 289  
Db 241 GICELHCPALVTYNTDFESMPNEGRTYTGASCVTACPNYISTDVG 289

RESULT 3  
US-08-422-734-1  
Sequence 1, Application US/08422734  
Patent No. 633169  
GENERAL INFORMATION:  
APPLICANT: Hudziak, Robert M.  
APPLICANT: Shepard, H. Michael  
APPLICANT: Ullrich, Axel  
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,734  
FILING DATE: 13-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/422108  
FILING DATE: 14-Apr-1995  
APPLICATION NUMBER: 08/355460  
FILING DATE: 13-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/048346  
FILING DATE: 15-APR-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/354319  
FILING DATE: 19-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 554C2D1  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-422-734-1

Query Match 100.0%; Score 1587; DB 4; Length 624;  
Best Local Similarity 100.0%; Pred. No. 6e-139;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVCTGDMKRLRPASPTHLDMLRHLRYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60  
|  
DB 1 STVCTGDMKRLRPASPTHLDMLRHLRYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60  
QY 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLREIQ 120  
|  
DB 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLREIQ 120  
QY 121 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDITNSRACHPCSPCKG 180  
|  
DB 121 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDITNSRACHPCSPCKG 180  
QY 181 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLHFHNS 240  
|  
DB 181 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLHFHNS 240  
QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289  
|  
DB 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289

## RESULT 4

US-09-146-283-4  
: Sequence 4, Application US/09146283  
: Patent NO. 5976546  
: GENERAL INFORMATION:  
: APPLICANT: Laus, Reiner  
: APPLICANT: Rugg, Curtis L.  
: APPLICANT: Wu, Hongyu  
: TITLE OF INVENTION: Immunostimulatory Compositions  
: NUMBER OF SEQUENCES: 10  
: CORRESPONDENCE ADDRESS:  
: ADDRESS: Dehlinger & Associates  
: STREET: 350 Cambridge Ave. Suite 250  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94306  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/146, 283  
: FILING DATE: 03-SEPT-1998  
: CLASSIFICATION: 536  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Judge, Linda R.  
: REGISTRATION NUMBER: 42,702  
: REFERENCE/DOCKET NUMBER: 7636-0010.21  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 650-324-0880  
: TELEFAX: 650-324-0960  
: INFORMATION FOR SEQ ID NO: 4:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 782 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: HYPOTHEICAL: NO  
: ORIGINAL SOURCE:  
: ORGANISM: homo sapiens  
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
US-09-146-283-4

## Query Match

Best Local Similarity 100.0%; Score 1587; DB 2; Length 782;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 STVCTGDMKRLRPASPTHLDMLRHLRYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60  
|  
DB 1 STVCTGDMKRLRPASPTHLDMLRHLRYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60

DB 22 STVCTGDMKRLRPASPTHLDMLRHLRYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 81  
QY 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLREIQ 120  
|  
DB 82 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLREIQ 141  
QY 121 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDITNSRACHPCSPCKG 180  
|  
DB 142 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDITNSRACHPCSPCKG 201  
QY 181 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLHFHNS 240  
|  
DB 202 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLHFHNS 261  
QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289  
|  
DB 262 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 310

## RESULT 5

US-08-579-823A-4  
: Sequence 4, Application US/08579823A  
: Patent NO. 6080409  
: GENERAL INFORMATION:  
: APPLICANT: Laus, Reiner  
: APPLICANT: Rugg, Curtis L.  
: APPLICANT: Wu, Hongyu  
: TITLE OF INVENTION: Immunostimulatory Composition and Method  
: NUMBER OF SEQUENCES: 10  
: CORRESPONDENCE ADDRESS:  
: ADDRESS: Dehlinger & Associates  
: STREET: 350 Cambridge Ave. Suite 250  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94306  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/579, 823A  
: FILING DATE: 03-DEC-1998  
: CLASSIFICATION: 536  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Judge, Linda R.  
: REGISTRATION NUMBER: 42,702  
: REFERENCE/DOCKET NUMBER: 7636-0010  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 650-324-0880  
: TELEFAX: 650-324-0960  
: INFORMATION FOR SEQ ID NO: 4:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 782 amino acids  
: TYPE: amino acid  
: TOPOLOGY: linear  
: MOLECULE TYPE: protein  
: HYPOTHEICAL: NO  
: ORIGINAL SOURCE:  
: ORGANISM: homo sapiens  
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
US-08-579-823A-4

## Query Match

Best Local Similarity 100.0%; Score 1587; DB 3; Length 782;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 STVCTGDMKRLRPASPTHLDMLRHLRYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60  
|  
DB 22 STVCTGDMKRLRPASPTHLDMLRHLRYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 81  
|  
QY 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLREIQ 120

Db 82 GYVLIHNOVROVPLORLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 141  
QY 121 LRSLETLKGVLIORNPOLCYODTILMKDIFPKNNOLATLIDTNSRACHCSPMCKG 180  
Db 142 LRSLETLKGVLIORNPOLCYODTILMKDIFPKNNOLATLIDTNSRACHCSPMCKG 201  
QY 181 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 240  
Db 202 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 261  
QY 241 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACPYNYLSTDVGS 289  
Db 262 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACPYNYLSTDVGS 310

## RESULT 6

US-09-344-195-4  
Sequence 4, Application US/09344195  
Patent No. 6210662  
GENERAL INFORMATION:  
APPLICANT: Laus, Reiner  
Ruegg, Curtis L.  
Mu, Hongyu  
TITLE OF INVENTION: Immunostimulatory Compositions  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave. Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/344,195  
FILING DATE: 24-Jun-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/146,283  
FILING DATE: 03-SEPT-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Judge, Linda R.  
REGISTRATION NUMBER: 42,702  
REFERENCE/DOCKET NUMBER: 7636-0010.21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 782 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: homo sapiens  
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-344-195-4

Query Match 100.0%; Score 1587; DB 4; Length 782;  
Best Local Similarity 100.0%; Pred. No. 8.1e-139;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCGVVGNNLELTLYLPTNASTSFLODIOEQV 60  
Db 22 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCGVVGNNLELTLYLPTNASTSFLODIOEQV 81

QY 61 GYVLIHNOVROVPLORLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 120  
Db 82 GYVLIHNOVROVPLORLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 141  
QY 121 LRSLETLKGVLIORNPOLCYODTILMKDIFPKNNOLATLIDTNSRACHCSPMCKG 180  
Db 142 LRSLETLKGVLIORNPOLCYODTILMKDIFPKNNOLATLIDTNSRACHCSPMCKG 201  
QY 181 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 240  
Db 202 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 261  
QY 241 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACPYNYLSTDVGS 289  
Db 262 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACPYNYLSTDVGS 310

## RESULT 7

US-08-467-083-68  
Sequence 68, Application US/08467083  
Patent No. 5726023  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Seed and Berry  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,083  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/414,417  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836 SEEDANBERY  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-467-083-68

Query Match 100.0%; Score 1587; DB 1; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.5e-138;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCGVVGNNLELTLYLPTNASTSFLODIOEQV 60  
Db 22 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCGVVGNNLELTLYLPTNASTSFLODIOEQV 81  
QY 61 GYVLIHNOVROVPLORLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 120  
Db 82 GYVLIHNOVROVPLORLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 141



Db 82 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLREIQ 141  
QY 121 LRSITELKGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
Db 142 LRSITELKGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201  
QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGRLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 240  
Db 202 SRCWGESSEDCQSLTRTVACGAGCARCKGRLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 261  
QY 241 GICELHCPALVYNTDTFESMPNEGRYTFGASCVTACPYNYLSTDVGS 289  
Db 262 GICELHCPALVYNTDTFESMPNEGRYTFGASCVTACPYNYLSTDVGS 310

## RESULT 10

US-08-486-348A-66  
Sequence 68, Application US/08486348A  
Patent No. 5846538  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,348A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-486-348A-66

Query Match 100.0%; Score 1587; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.5e-138;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STGYCTGDMKLRPLASPEHLDMLRHLVGGCOVVGNGLELYLPTNASISFLQDDIOEQ 60  
Db 22 STGYCTGDMKLRPLASPEHLDMLRHLVGGCOVVGNGLELYLPTNASISFLQDDIOEQ 81  
QY 61 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLREIQ 120  
Db 82 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLREIQ 141  
QY 121 LRSITELKGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
Db 142 LRSITELKGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201

QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGRLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 240  
Db 202 SRCWGESSEDCQSLTRTVACGAGCARCKGRLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 261  
QY 241 GICELHCPALVYNTDTFESMPNEGRYTFGASCVTACPYNYLSTDVGS 289  
Db 262 GICELHCPALVYNTDTFESMPNEGRYTFGASCVTACPYNYLSTDVGS 310

## RESULT 11

US-08-625-101-2  
Sequence 2, Application US/08625101  
Patent No. 5869445  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
TITLE OF INVENTION: ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/625,101  
FILING DATE: 01-APR-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-625-101-2

Query Match 100.0%; Score 1587; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.5e-138;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STGYCTGDMKLRPLASPEHLDMLRHLVGGCOVVGNGLELYLPTNASISFLQDDIOEQ 60  
Db 22 STGYCTGDMKLRPLASPEHLDMLRHLVGGCOVVGNGLELYLPTNASISFLQDDIOEQ 81  
QY 61 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLREIQ 120  
Db 82 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLREIQ 141  
QY 121 LRSITELKGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
Db 142 LRSITELKGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201  
QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGRLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 240  
Db 202 SRCWGESSEDCQSLTRTVACGAGCARCKGRLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 261  
QY 241 GICELHCPALVYNTDTFESMPNEGRYTFGASCVTACPYNYLSTDVGS 289



Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

## RESULT 12

US-08-468-545B-68  
; Sequence 68, Application US/08468545B  
; Patent No. 5876712  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,545B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 68:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-468-545B-68

Query Match 100.0%; Score 1587; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.5e-138;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNASLSFLDIOEVO 60  
Db 22 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNASLSFLDIOEVO 81  
QY 61 GYVLIANQVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 120  
Db 82 GYVLIANQVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 141  
QY 121 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180  
Db 142 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 201  
QY 181 SRCWESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 240  
Db 202 SRCWESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 261  
QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289  
Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 13  
US-08-356-786-2  
; Sequence 2, Application US/08356786

; Patent No. 5877305  
; GENERAL INFORMATION:  
; APPLICANT: Huston, James S.  
; APPLICANT: Oppermann, Hermann  
; APPLICANT: Houston, L. L.  
; APPLICANT: Ring, David B.  
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
; TITLE OF INVENTION: Marker  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
; STREET: Exchange Place, 53 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/356,786  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/831,967  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-053  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7100  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1255 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-356-786-2

Query Match 100.0%; Score 1587; DB 2; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.5e-138;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNASLSFLDIOEVO 60  
Db 22 STVCTGDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNASLSFLDIOEVO 81  
QY 61 GYVLIANQVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 120  
Db 82 GYVLIANQVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 141  
QY 121 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180  
Db 142 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 201  
QY 181 SRCWESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 240  
Db 202 SRCWESSEDCQSLRTVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNS 261  
QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289  
Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 14  
US-08-466-680B-68  
; Sequence 68, Application US/08466680B  
; Patent No. 6073122  
; GENERAL INFORMATION:

```

: APPLICANT: Cheever, Martin A.
: APPLICANT: Disis, Mary L.
: TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
: TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
: TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
: NUMBER OF SEQUENCES: 69
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seed and Berry LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/466,680B
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ. ID NO: 68:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
US-08-466-680B-68

Query Match 100.0%; Score 1587; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-138;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVCTGCTDMKRLRPA SPETHLDMRLHLYGCGVVGNIETLYLPTNASSLFLDIOIEVQ 60
DB 22 STVCTGCTDMKRLRPA SPETHLDMRLHLYGCGVVGNIETLYLPTNASSLFLDIOIEVQ 81
QY 61 GYVLIANQVQVPLQRLRIYRGTOLEEDNYALAVLDNGDPLNNTTPTVTGASPGGLREIQ 120
DB 82 GYVLIANQVQVPLQRLRIYRGTOLEEDNYALAVLDNGDPLNNTTPTVTGASPGGLREIQ 141
QY 121 LRSLETILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSRACHCSPCKG 180
DB 142 LRSLETILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSRACHCSPCKG 201
QY 181 SRCMGESSEDCQSLTRTYVACGACGRCGKPLPTCCHEQCAAGCTGPKHSDCLAHFNHS 240
DB 202 SRCMGESSEDCQSLTRTYVACGACGRCGKPLPTCCHEQCAAGCTGPKHSDCLAHFNHS 261
QY 241 GICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPYNYLSTDVGS 289
DB 262 GICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPYNYLSTDVGS 310

RESULT 15
US-08-336-708A-9
: Sequence 9, Application US/08336708A
: Patent No. 5521295
: GENERAL INFORMATION:
: APPLICANT: Pacific, Robert E.
: APPLICANT: Thomson, Arlen R.
: APPLICANT: Chang, Ming-Shi
: TITLE OF INVENTION: Hybrid Receptor Molecules
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Amgen Inc.

```

```

: STREET: 1840 Dehavilland Drive
: CITY: Thousand Oaks
: STATE: California
: COUNTRY: USA
: ZIP: 91320-1789
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/336,708A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Oleski, Nancy
: REFERENCE/DOCKET NUMBER: A-241A
: INFORMATION FOR SEQ. ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 644 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-336-708A-9

Query Match 45.1%; Score 716.5; DB 1; Length 644;
Best Local Similarity 46.5%; Pred. No. 2.9e-58;
Matches 134; Conservative 44; Mismatches 99; Indels 11; Gaps 2;

QY 3 QVCTGCTDMKRLRPA SPETHLDMRLHLYGCGVVGNIETLYLPTNASSLFLDIOIEVQ 62
DB 29 KVCQGSTNKLTLQGLTFEDHFLSLQRMFNCEVVLGNLEITYVORNDLSFLKTIQIEVAGY 88
QY 63 VLIANQVQVPLQRLRIYRGTOLEEDNYALAVLDNGDPLNNTTPTVTGASPGGLREIQ 122
DB 89 VLIANQVQVPLQRLRIYRGTOLEEDNYALAVLDNGDPLNNTTPTVTGASPGGLREIQ 138
QY 123 SLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSRACHCSPCKGSR 182
DB 139 NLGELLHGAVERFSNNPALCNVESIQWRDIYSDPFLSNMSKDFONHLGSCCKDPCSPGNS 198
QY 183 CMGESSSEDCQSLTRTYVACGACGCA-RCKGRLPTDCHEQCAAGCTGPKHSDCLAHFNHS 241
DB 199 CMGAGEENCCOKLTKIICAOQCSGRCKRSPSDCHNCAAGCTGPKHSDCLAHFNHS 258
QY 242 ICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPYNYLSTDVGS 289
DB 259 TCKDTCPPLMLNPTTYQMDVNPDEGRYTFGATCVKCPKRYVYTDHGS 306

Search completed: April 28, 2003, 13:43:18
Job time : 15.4574 secs

```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:40:24 ; Search time 10.6321 Seconds  
(without alignments)

2178.074 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STGVCTGDMKRLRPASPER.....FGASCVTACPYNTLSDVGS 289

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1587	100.0	289	10 US-09-821-883-23	Sequence 23, Appl
2	1587	100.0	479	10 US-09-821-883-5	Sequence 5, Appl
3	1587	100.0	555	10 US-09-821-883-1	Sequence 1, Appl
4	1587	100.0	564	10 US-09-821-883-3	Sequence 3, Appl
5	1587	100.0	645	10 US-09-821-161-1	Sequence 1, Appl
6	1587	100.0	653	9 US-09-854-356-3	Sequence 3, Appl
7	1587	100.0	690	10 US-09-821-883-2	Sequence 2, Appl
8	1587	100.0	697	10 US-09-821-883-4	Sequence 4, Appl
9	1587	100.0	712	9 US-09-854-356-7	Sequence 7, Appl
10	1587	100.0	919	9 US-09-854-356-6	Sequence 6, Appl
11	1587	100.0	1255	9 US-09-769-508-2	Sequence 2, Appl
12	1587	100.0	1255	9 US-09-854-356-1	Sequence 1, Appl
13	1587	100.0	1255	9 US-09-930-125-2	Sequence 2, Appl
14	1587	100.0	1255	9 US-09-441-411-6	Sequence 6, Appl
15	1587	100.0	1255	10 US-09-811-123-9	Sequence 9, Appl
16	1587	100.0	1255	10 US-09-811-115-3	Sequence 3, Appl
17	1353.5	85.3	654	9 US-09-854-356-8	Sequence 8, Appl
18	1353.5	85.3	1256	9 US-09-854-356-2	Sequence 2, Appl
19	1353.5	85.3	1260	9 US-09-870-759-118	Sequence 118, App

20	1352.5	85.2	1256	9 US-09-854-356-14	Sequence 14, Appl
21	878	55.3	191	9 US-09-441-411-9	Sequence 9, Appl
22	716.5	45.1	657	9 US-10-172-620-18	Sequence 18, Appl
23	716.5	45.1	1210	10 US-09-725-433-2	Sequence 2, Appl
24	716.5	45.1	1308	10 US-09-940-101-2	Sequence 2, Appl
25	713.5	45.0	615	10 US-09-940-101-4	Sequence 4, Appl
26	700.5	44.1	478	10 US-09-867-521-2	Sequence 2, Appl
27	686	43.2	1342	9 US-10-172-620-16	Sequence 16, Appl
28	249	15.7	120	9 US-10-172-620-17	Sequence 17, Appl
29	232.5	14.7	1367	9 US-09-870-759-120	Sequence 120, App
30	186	11.7	366	10 US-09-205-658-103	Sequence 103, App
31	186	11.7	366	10 US-09-844-353A-103	Sequence 103, App
32	185	11.7	1724	10 US-09-205-658-12	Sequence 12, Appl
33	185	11.7	1724	10 US-09-844-353A-12	Sequence 12, Appl
34	171.5	10.8	383	10 US-09-205-658-105	Sequence 105, App
35	171.5	10.8	383	10 US-09-844-353A-105	Sequence 105, App
36	166.5	10.5	370	10 US-09-205-658-104	Sequence 104, App
37	166.5	10.5	370	10 US-09-844-353A-104	Sequence 104, App
38	132	8.3	381	10 US-09-205-658-106	Sequence 106, App
39	132	8.3	381	10 US-09-844-353A-106	Sequence 106, App
40	117	7.4	2290	9 US-10-123-155-123	Sequence 323, App
41	116.5	7.3	68	10 US-09-466-320-11	Sequence 11, Appl
42	116	7.3	22	10 US-09-466-320-19	Sequence 19, Appl
43	115.5	7.3	1875	9 US-10-123-155-359	Sequence 359, App
44	114.5	7.2	1281	9 US-10-123-155-509	Sequence 509, App
45	114	7.2	840	9 US-10-123-155-227	Sequence 227, App

## ALIGNMENTS

```
RESULT 1
US-09-821-883-23
Sequence 23, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Relner
APPLICANT: Vidovic, Thomas
TITLE OF INVENTION: gradids, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT FILING DATE: 2001-03-30
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 289
TYPE: PRT
ORGANISM: Homo sapiens
US-09-821-883-23

Query Match 100.0%; Score 1587; DB 10; Length 289;
Best Local Similarity 100.0%; Pred. No. 6.2e-120;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STGVCTGDMKRLRPASPERHLMRLHYGCGVQVGMLELTLYPTNLSLFLDIOEVQ 60
DB 1 STGVCTGDMKRLRPASPERHLMRLHYGCGVQVGMLELTLYPTNLSLFLDIOEVQ 60
QY 61 GYVLIHQVQVPLQRIKRIYRGVQLFEDNYALAVLNDGDPPLNNTPTVTSGLRELQ 120
DB 61 GYVLIHQVQVPLQRIKRIYRGVQLFEDNYALAVLNDGDPPLNNTPTVTSGLRELQ 120
QY 121 LRLSLLEIKGVLJORNQOLCYODTILMKDIFHKNNQALATLIDTNRACHPSCPMKG 180
DB 121 LRLSLLEIKGVLJORNQOLCYODTILMKDIFHKNNQALATLIDTNRACHPSCPMKG 180
QY 181 SRCWESSEDCOSLTRIVCAGGACARCKGPLPTDCHEOCAGCTGPKHSXCLACHFNHS 240
DB 181 SRCWESSEDCOSLTRIVCAGGACARCKGPLPTDCHEOCAGCTGPKHSXCLACHFNHS 240
```

OY 241 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTACPNYLISTDVG 289  
 Db 241 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTACPNYLISTDVG 289

# RESULT 2

US-09-821-883-5  
 ; Sequence 5, Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Graddis, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic  
 ; FILE REFERENCE: 7636-0022.30  
 ; CURRENT APPLICATION NUMBER: US/09/821,883  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR FILING DATE: 2000-03-30  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 479  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HER300\*-rgm-CSF construct  
 US-09-821-883-5

Query Match 100.0%; Score 1587; DB 10; Length 479;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STGVCTGDMKLRIPASPEHLDMLRHLVGGCGVGNLELTPLTNASLSFLQDIOEVQ 60  
 Db 41 STGVCTGDMKLRIPASPEHLDMLRHLVGGCGVGNLELTPLTNASLSFLQDIOEVQ 100  
 OY 61 GYVLIHNOVROYPLQRLIRVGTQLEFDNYALAVLDNGDPLNNTTPVTGASGGLREIQ 120  
 Db 101 GYVLIHNOVROYPLQRLIRVGTQLEFDNYALAVLDNGDPLNNTTPVTGASGGLREIQ 160  
 OY 121 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHCSPMCKG 180  
 Db 161 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHCSPMCKG 220  
 OY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLHFNHS 240  
 Db 221 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLHFNHS 280  
 OY 241 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTACPNYLISTDVG 289  
 Db 281 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTACPNYLISTDVG 329

# RESULT 3

US-09-821-883-1  
 ; Sequence 1, Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Graddis, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic  
 ; FILE REFERENCE: 7636-0022.30  
 ; CURRENT APPLICATION NUMBER: US/09/821,883  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR FILING DATE: 2000-03-30  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1  
 ; LENGTH: 555  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HER500 construct  
 US-09-821-883-1

Query Match 100.0%; Score 1587; DB 10; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STGVCTGDMKLRIPASPEHLDMLRHLVGGCGVGNLELTPLTNASLSFLQDIOEVQ 60  
 Db 41 STGVCTGDMKLRIPASPEHLDMLRHLVGGCGVGNLELTPLTNASLSFLQDIOEVQ 100  
 OY 61 GYVLIHNOVROYPLQRLIRVGTQLEFDNYALAVLDNGDPLNNTTPVTGASGGLREIQ 120  
 Db 101 GYVLIHNOVROYPLQRLIRVGTQLEFDNYALAVLDNGDPLNNTTPVTGASGGLREIQ 160  
 OY 121 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHCSPMCKG 180  
 Db 161 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHCSPMCKG 220  
 OY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLHFNHS 240  
 Db 221 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLHFNHS 280  
 OY 241 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTACPNYLISTDVG 289  
 Db 281 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTACPNYLISTDVG 329

# RESULT 4

US-09-821-883-3  
 ; Sequence 3, Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Graddis, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic  
 ; FILE REFERENCE: 7636-0022.30  
 ; CURRENT APPLICATION NUMBER: US/09/821,883  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR FILING DATE: 2000-03-30  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 564  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HER500\* construct  
 US-09-821-883-3

Query Match 100.0%; Score 1587; DB 10; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STGVCTGDMKLRIPASPEHLDMLRHLVGGCGVGNLELTPLTNASLSFLQDIOEVQ 60  
 Db 41 STGVCTGDMKLRIPASPEHLDMLRHLVGGCGVGNLELTPLTNASLSFLQDIOEVQ 100  
 OY 61 GYVLIHNOVROYPLQRLIRVGTQLEFDNYALAVLDNGDPLNNTTPVTGASGGLREIQ 120  
 Db 101 GYVLIHNOVROYPLQRLIRVGTQLEFDNYALAVLDNGDPLNNTTPVTGASGGLREIQ 160  
 OY 121 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHCSPMCKG 180  
 Db 161 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHCSPMCKG 220

QY 181 SRCWSESDCOSLTRVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 240  
 |||  
 Db 221 SRCWSESDCOSLTRVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 280  
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289  
 |||  
 Db 281 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329

## RESULT 5

US-09-921-161-1  
 : Sequence 1, Application US/09921161  
 : Patent No. US20020090662A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Ralph, Peter  
 : TITLE OF INVENTION: ANALYTICAL METHOD  
 : FILE REFERENCE: GENENT.066A  
 : CURRENT APPLICATION NUMBER: US/09/921,161  
 : CURRENT FILING DATE: 2001-08-01  
 : PRIOR APPLICATION NUMBER: 60/225,433  
 : PRIOR FILING DATE: 2000-08-15  
 : NUMBER OF SEQ ID NOS: 1  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 1  
 : LENGTH: 645  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 : US-09-921-161-1

Query Match 100.0%; Score 1587; DB 10; Length 645;  
 Best Local Similarity 100.0%; Pred. No. 1,6e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKRLRPASETHLDMRLHYOCQVQVGNLELTYPNTASLSFLDIOEYQ 60  
 |||  
 Db 22 STQVCTGDMKRLRPASETHLDMRLHYOCQVQVGNLELTYPNTASLSFLDIOEYQ 81  
 QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120  
 |||  
 Db 82 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 141  
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIHKNQALATLIDTNRSRACHPCSPMKG 180  
 |||  
 Db 142 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIHKNQALATLIDTNRSRACHPCSPMKG 201  
 QY 181 SRCWSESDCOSLTRVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 240  
 |||  
 Db 202 SRCWSESDCOSLTRVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 261  
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289  
 |||  
 Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

## RESULT 6

US-09-854-356-3  
 : Sequence 3, Application US/09854356  
 : Patent No. US2002017367A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Cheever, Martin A.  
 : APPLICANT: Gheysen, Dirk  
 : APPLICANT: Corixa Corporation  
 : TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 : FILE REFERENCE: 014058-009810PC  
 : CURRENT APPLICATION NUMBER: US/09/854,356  
 : CURRENT FILING DATE: 2001-05-09  
 : PRIOR APPLICATION NUMBER: US 09/493,480  
 : PRIOR FILING DATE: 2000-01-28  
 : PRIOR APPLICATION NUMBER: US 60/117,976  
 : PRIOR FILING DATE: 1999-01-29  
 : NUMBER OF SEQ ID NOS: 26

: SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 3  
 : LENGTH: 653  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu  
 : US-09-854-356-3

Query Match 100.0%; Score 1587; DB 9; Length 653;  
 Best Local Similarity 100.0%; Pred. No. 1,6e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKRLRPASETHLDMRLHYOCQVQVGNLELTYPNTASLSFLDIOEYQ 60  
 |||  
 Db 22 STQVCTGDMKRLRPASETHLDMRLHYOCQVQVGNLELTYPNTASLSFLDIOEYQ 81  
 QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120  
 |||  
 Db 82 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 141  
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIHKNQALATLIDTNRSRACHPCSPMKG 180  
 |||  
 Db 142 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIHKNQALATLIDTNRSRACHPCSPMKG 201  
 QY 181 SRCWSESDCOSLTRVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 240  
 |||  
 Db 202 SRCWSESDCOSLTRVYACAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 261  
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289  
 |||  
 Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

## RESULT 7

US-09-821-883-2  
 : Sequence 2, Application US/09821883  
 : Patent No. US20020061310A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Laus, Reiner  
 : APPLICANT: Vidovic, Damlir  
 : TITLE OF INVENTION: Compositions and Methods for Dendritic  
 : FILE REFERENCE: 7636-0022.30  
 : CURRENT APPLICATION NUMBER: US/09/821,883  
 : CURRENT FILING DATE: 2001-03-30  
 : PRIOR APPLICATION NUMBER: US 60/193,504  
 : PRIOR FILING DATE: 2000-03-30  
 : NUMBER OF SEQ ID NOS: 30  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 2  
 : LENGTH: 690  
 : TYPE: PRT  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: HER500-hgm-CSF construct  
 : US-09-821-883-2

Query Match 100.0%; Score 1587; DB 10; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 1,7e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKRLRPASETHLDMRLHYOCQVQVGNLELTYPNTASLSFLDIOEYQ 60  
 |||  
 Db 41 STQVCTGDMKRLRPASETHLDMRLHYOCQVQVGNLELTYPNTASLSFLDIOEYQ 100  
 QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120  
 |||  
 Db 101 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 160  
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIHKNQALATLIDTNRSRACHPCSPMKG 180  
 |||

Db 161 LRSLETILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 220  
 QY 181 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 240  
 Db 221 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 280  
 QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTFGASCVCAPYNYLSTDVGS 289  
 Db 281 GICELHCPALVYNTDTFESMPNPEGRTYTFGASCVCAPYNYLSTDVGS 329

RESULT 8  
 US-09-821-883-4  
 ; Sequence 4, Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Graddis, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic  
 ; FILE OF INVENTION: Cell-Based Immunotherapy  
 ; FILE REFERENCE: 7636-0022.30  
 ; CURRENT APPLICATION NUMBER: US/09/821,883  
 ; CURRENT FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: US 60/193,504  
 ; PRIOR FILING DATE: 2000-03-30  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 697  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HER500+-igm-CSF construct  
 US-09-821-883-4

Query Match 100.0%; Score 1587; DB 10; Length 697;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDMKRLRPASPETHLDMRLHYGCGVQVGNLELYLPTNASTLFLQDIOEVQ 60  
 Db 41 STQVCTGTDMKRLRPASPETHLDMRLHYGCGVQVGNLELYLPTNASTLFLQDIOEVQ 100  
 QY 61 GYVLIHNOVROYPLQRLIRVSTQLEFDNYALAVLDNGDPLNNTTPVYGASPGGLREIQ 120  
 Db 101 GYVLIHNOVROYPLQRLIRVSTQLEFDNYALAVLDNGDPLNNTTPVYGASPGGLREIQ 160  
 QY 121 LRSLETILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
 Db 161 LRSLETILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 220  
 QY 181 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 240  
 Db 221 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 280  
 QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTFGASCVCAPYNYLSTDVGS 289  
 Db 281 GICELHCPALVYNTDTFESMPNPEGRTYTFGASCVCAPYNYLSTDVGS 329

RESULT 9  
 US-09-854-356-7  
 ; Sequence 7, Application US/09854356  
 ; Patent No. US20020177567A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Cheysen, Dirk  
 ; APPLICANT: Corixa Corporation  
 ; APPLICANT: SmithKline Beecham Biologicals S. A.  
 ; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 ; FILE REFERENCE: 014058-009810PC  
 ; CURRENT APPLICATION NUMBER: US/09/854,356

; CURRENT FILING DATE: 2001-05-09  
 ; PRIOR APPLICATION NUMBER: US 09/493,480  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: US 60/117,976  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 712  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 ; OTHER INFORMATION: of ECD and delta PD of human HER-2/neu  
 US-09-854-356-7

Query Match 100.0%; Score 1587; DB 9; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDMKRLRPASPETHLDMRLHYGCGVQVGNLELYLPTNASTLFLQDIOEVQ 60  
 Db 22 STQVCTGTDMKRLRPASPETHLDMRLHYGCGVQVGNLELYLPTNASTLFLQDIOEVQ 81  
 QY 61 GYVLIHNOVROYPLQRLIRVSTQLEFDNYALAVLDNGDPLNNTTPVYGASPGGLREIQ 120  
 Db 82 GYVLIHNOVROYPLQRLIRVSTQLEFDNYALAVLDNGDPLNNTTPVYGASPGGLREIQ 141  
 QY 121 LRSLETILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180  
 Db 142 LRSLETILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201  
 QY 181 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 240  
 Db 202 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEQCAAGCTGPKHSDCLACLFHNHS 261  
 QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTFGASCVCAPYNYLSTDVGS 289  
 Db 262 GICELHCPALVYNTDTFESMPNPEGRTYTFGASCVCAPYNYLSTDVGS 310

RESULT 10  
 US-09-854-356-6  
 ; Sequence 6, Application US/09854356  
 ; Patent No. US20020177567A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Cheysen, Dirk  
 ; APPLICANT: Corixa Corporation  
 ; APPLICANT: SmithKline Beecham Biologicals S. A.  
 ; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 ; FILE REFERENCE: 014058-009810PC  
 ; CURRENT APPLICATION NUMBER: US/09/854,356  
 ; CURRENT FILING DATE: 2001-05-09  
 ; PRIOR APPLICATION NUMBER: US 09/493,480  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: US 60/117,976  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 919  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 ; OTHER INFORMATION: of ECD and PD of human HER-2/neu  
 US-09-854-356-6

Query Match 100.0%; Score 1587; DB 9; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-119;  
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 STOVCTGDMKRLRPASBETHLDMRLHLYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 60
D 22 STOVCTGDMKRLRPASBETHLDMRLHLYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 81
QY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 120
D 82 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 180
D 142 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 201
QY 181 SRCWESSEDCQSLRTVACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
D 202 SRCWESSEDCQSLRTVACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCAALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 289
D 262 GICELHCAALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 310

```

# RESULT 11

```

US-09-769-508-2
; Sequence 2, Application US/09769508
; Patent No. US20020155527A1
; GENERAL INFORMATION:
; APPLICANT: STUART, SUSAN G.
; APPLICANT: MONAHAN, JOHN J.
; APPLICANT: LANGTON, BEATRICE CLAUDIA
; APPLICANT: HANCOCK, MIRIAM E.C.
; APPLICANT: CHAO, LORRINE A.
; APPLICANT: BLUFORD, PETER
; TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
; FILE REFERENCE: BBIO-111-C1
; CURRENT APPLICATION NUMBER: US/09/769,508
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO: 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-508-2

```

```

Query Match 100.0%; Score 1587; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3,4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STOVCTGDMKRLRPASBETHLDMRLHLYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 60
D 22 STOVCTGDMKRLRPASBETHLDMRLHLYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 81
QY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 120
D 82 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 180
D 142 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 201
QY 181 SRCWESSEDCQSLRTVACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
D 202 SRCWESSEDCQSLRTVACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCAALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 289
D 262 GICELHCAALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 310

```

```

RESULT 12
US-09-854-356-1
; Sequence 1, Application US/09854356
; Patent No. US2002017567A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Cheyssen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO: 1
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(653)
; OTHER INFORMATION: extracellular domain (ECD)
; NAME/KEY: DOMAIN
; LOCATION: (676)..(1255)
; OTHER INFORMATION: intracellular domain (ICD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1255)
; OTHER INFORMATION: phosphorylation domain (PD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1048)
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred
; OTHER INFORMATION: portion (delta PD)
US-09-854-356-1

```

```

Query Match 100.0%; Score 1587; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3,4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STOVCTGDMKRLRPASBETHLDMRLHLYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 60
D 22 STOVCTGDMKRLRPASBETHLDMRLHLYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 81
QY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 120
D 82 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 180
D 142 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 201
QY 181 SRCWESSEDCQSLRTVACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
D 202 SRCWESSEDCQSLRTVACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCAALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 289
D 262 GICELHCAALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 310

```

# RESULT 13

```

US-09-930-125-2
; Sequence 2, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.

```

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
;; FILE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
;; FILE REFERENCE: 210121.544  
;; CURRENT APPLICATION NUMBER: US/09/930.125  
;; CURRENT FILING DATE: 2001-08-14  
;; NUMBER OF SEQ ID NOS: 25  
;; SOFTWARE: FASTSEQ for Windows Version 3.0  
;; SEQ ID NO 2  
;; LENGTH: 1255  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-930-125-2

Query Match 100.0%; Score 1587; DB 9; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 3.4e-119;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STOVCTGCTDMKRLRLPASPEHLDMLRLHYOGCOVVOGNLELYLPTNASISFLDIOIEVQ 60  
DB 22 STOVCTGCTDMKRLRLPASPEHLDMLRLHYOGCOVVOGNLELYLPTNASISFLDIOIEVQ 81  
OY 61 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120  
DB 82 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141  
OY 121 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKG 180  
DB 142 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKG 201  
OY 181 SRCWGESSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPHSDCLACLHFNHS 240  
DB 202 SRCWGESSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPHSDCLACLHFNHS 261  
OY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYLISTDVGS 289  
DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYLISTDVGS 310

RESULT 14  
US-09-441-411-6  
; Sequence 6, Application US/09441411  
; Publication No. US20030008342A1  
; GENERAL INFORMATION:

; APPLICANT: Scholler, Nathalie B.  
; APPLICANT: Disis, Mary L.  
; APPLICANT: Helstrom, Ingegerd  
; APPLICANT: Helstrom, Karl Erik  
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES  
; FILE REFERENCE: 730033.409  
; CURRENT APPLICATION NUMBER: US/09/441.411  
; CURRENT FILING DATE: 1999-11-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-441-411-6

Query Match 100.0%; Score 1587; DB 9; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 3.4e-119;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STOVCTGCTDMKRLRLPASPEHLDMLRLHYOGCOVVOGNLELYLPTNASISFLDIOIEVQ 60  
DB 22 STOVCTGCTDMKRLRLPASPEHLDMLRLHYOGCOVVOGNLELYLPTNASISFLDIOIEVQ 81  
OY 61 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120  
DB 82 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141  
OY 121 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKG 180

DB 142 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKG 201  
OY 181 SRCWGESSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPHSDCLACLHFNHS 240  
DB 202 SRCWGESSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPHSDCLACLHFNHS 261  
OY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYLISTDVGS 289  
DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYLISTDVGS 310

RESULT 15  
US-09-811-123-9  
; Sequence 9, Application US/09811123  
; Patent No. US2002001587A1  
; GENERAL INFORMATION:

; APPLICANT: Sharon Erickson  
; APPLICANT: Ralph Schwall  
; APPLICANT: Mark Sliwowski  
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERB  
; FILE REFERENCE: GENENT.073A2  
; CURRENT APPLICATION NUMBER: US/09/811.123  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/238.327  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 09/602.530  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-811-123-9

Query Match 100.0%; Score 1587; DB 10; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 3.4e-119;  
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STOVCTGCTDMKRLRLPASPEHLDMLRLHYOGCOVVOGNLELYLPTNASISFLDIOIEVQ 60  
DB 22 STOVCTGCTDMKRLRLPASPEHLDMLRLHYOGCOVVOGNLELYLPTNASISFLDIOIEVQ 81  
OY 61 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120  
DB 82 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141  
OY 121 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKG 180  
DB 142 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKG 201  
OY 181 SRCWGESSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPHSDCLACLHFNHS 240  
DB 202 SRCWGESSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPHSDCLACLHFNHS 261  
OY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYLISTDVGS 289  
DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPNYLISTDVGS 310

Search completed: April 28, 2003, 13:44:07  
Job time: 14.6321 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:43:24 ; Search time 22 Seconds  
(Without-alignments)

3015.125 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 690

Sequence: 1 MRAAPLLARASLSGLF.....EPVQGAAPPAAHHHHH 690

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	292	42.3	1255	1 A24571	protein-tyrosine k
2	127	18.4	1444	1 F0H0GM	granulocyte-macrop
3	67	9.7	1234	1 I48161	p-185 precursor -
4	52	7.5	1260	1 TVRTNU	protein-tyrosine k
5	35	5.1	386	1 JH0610	acid phosphatase (
6	22	3.2	144	1 A61632	granulocyte-macrop
7	22	3.2	144	1 JH0469	granulocyte-macrop
8	11	1.6	144	2 A44936	granulocyte-macrop
9	10	1.4	143	1 F0B0GM	granulocyte-macrop
10	9	1.3	527	2 A42032	epidermal growth f
11	9	1.3	644	2 A36325	epidermal growth f
12	9	1.3	1210	1 G0HUE	epidermal growth f
13	9	1.3	1210	2 A53183	epidermal growth f
14	9	1.3	1223	1 TVCHRV	epidermal growth f
15	8	1.2	127	2 I46259	granulocyte-macrop
16	8	1.2	265	2 A46534	ubiquitin-cytochro
17	8	1.2	268	2 A56446	ig heavy chain v r
18	8	1.2	325	2 D70666	probable mod prot
19	8	1.2	341	2 B83298	conserved hypotet
20	8	1.2	348	2 D95067	phenylalanyl-trna
21	8	1.2	354	2 F87398	hyd family secret
22	8	1.2	355	2 T44850	molybdopterin co-f
23	8	1.2	374	2 I39637	molybdopterin co-f
24	8	1.2	375	2 C97935	phenylalanine-trna
25	8	1.2	463	2 H70922	hypothetical prote
26	8	1.2	517	2 T31463	probable diaphophyt
27	8	1.2	539	2 AH0508	sensor kinase cita
28	8	1.2	547	2 S70538	signal-transducing
29	8	1.2	583	2 D90052	hypothetical prote

30	8	1.2	610	2 T32327	hypothetical prote
31	8	1.2	715	2 S73637	ATP-dependent prot
32	8	1.2	825	1 EDBEXD	immediate-early pr
33	8	1.2	1851	2 T13980	calcium channel al
34	7	1.0	62	2 F90177	hypothetical prote
35	7	1.0	65	2 T29066	hypothetical prote
36	7	1.0	68	2 E69489	LSU ribosomal prot
37	7	1.0	82	2 T30135	hypothetical prote
38	7	1.0	89	2 A10221	flagellar biosynth
39	7	1.0	98	2 E90606	hypothetical prote
40	7	1.0	104	1 G69482	hypothetical prote
41	7	1.0	41	2 C35826	hypothetical 13k p
42	7	1.0	114	2 E72600	hypothetical prote
43	7	1.0	120	2 T35946	probable regulator
44	7	1.0	125	1 TCCHRP	calcitonin gene-re
45	7	1.0	136	1 TCON	calcitonin 1 precu

## ALIGNMENTS

RESULT 1  
A24571  
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human  
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e  
C:Species: Homo sapiens (man)  
C>Date: 25-Oct-1987 #sequence\_revision 06-Dec-1996 #text\_change 11-Jun-1999  
C:Accession: A24571; A23491; A44188; B44188; I59509; I57622  
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T  
Nature 319, 230-234, 1986  
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth  
A:Reference number: A24571; MUID:86118663; PMID:3003577  
A:Accession: A24571  
A:Molecule type: mRNA  
A:Residues: 1-1255 <YAM>  
A:Cross-references: GB:X03363; NID:931197; PIDN:CAA27060.1; PID:931198  
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985  
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid  
A:Reference number: A25491; MUID:86016729; PMID:2995967  
A:Accession: A25491  
A:Molecule type: DNA  
A:Residues: 737-1031 <SEM>  
A:Cross-references: GB:M1767; NID:g182163; PIDN:AAA35808.1; PID:g553282  
R:Consens. L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg  
Science 230, 1132-1139, 1985  
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro  
A:Reference number: A44188; MUID:86070181; PMID:2999974  
A:Accession: A44188  
A:Molecule type: DNA  
A:Residues: 740-910 <COU1>  
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989  
A:Accession: B44188  
A:Molecule type: mRNA  
A:Residues: 1-517, 'RAL', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>  
A:Cross-references: GB:M1730; NID:g183986  
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.  
Science 229, 974-976, 1985  
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.  
A:Reference number: I59509; MUID:85272597; PMID:2992089  
A:Accession: I59509  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 832-909 <REX>  
A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808  
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.  
Mol. Cell. Biol. 7, 2597-2601, 1987  
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio  
A:Reference number: I57622; MUID:87286898; PMID:3039351  
A:Accession: I57622  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-191 <TAU>

A:Cross-references: GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:g553332  
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
 C:Genetics:  
 A:Gene: GDB:ERBB2; NGL; NEU; HER-2  
 A:Cross-references: GDB:120613; OMIM:164870  
 A:Map position: 17q21.1-17q21.1  
 A:Introns: 25/1; 75/3; 147/1; 883/3  
 A:Note: The list of introns is incomplete  
 C:Function:  
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 lase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-155/Product: protein-tyrosine kinase erbB2 #status predicted <MAY>  
 F:22-653/Domain: extracellular #status predicted <EXT>  
 F:70-304/Domain: EGF receptor extracellular domain repeat <EB1>  
 F:395-605/Domain: EGF receptor extracellular domain repeat <EB2>  
 F:654-675/Domain: transmembrane #status predicted <TM>  
 F:676-1255/Domain: intracellular #status predicted <INT>  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif  
 F:68,124,187,259,530,571,629/Binding site: carboxydrate (Asn) (covalent) #status predicted  
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:753/Active site: Lys #status predicted  
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 42.3%; Score 292; DB 1; Length 1255;  
 Best Local Similarity 100.0%; Pred. No. 6,9e-281;  
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAATGTCGTGDMKLRPAPSETHLDMRLHLYGCGVYVGNLELTPTNALSFLDDIO 97  
 DB 19 GAATGTCGTGDMKLRPAPSETHLDMRLHLYGCGVYVGNLELTPTNALSFLDDIO 78  
 OY 98 EVGQYVLIANNOVROVPLQRLRYGTQLFEDNVALAVLDNGDPLNNTPTPTGSPGGLR 157  
 DB 79 EVGQYVLIANNOVROVPLQRLRYGTQLFEDNVALAVLDNGDPLNNTPTPTGSPGGLR 138  
 OY 158 ELQRLSTLEIKGVLQIQRNQLCYQDTILMKDIFHRNNOLATLIDNRSRACHPCSPM 217  
 DB 139 ELQRLSTLEIKGVLQIQRNQLCYQDTILMKDIFHRNNOLATLIDNRSRACHPCSPM 198  
 OY 218 CKGRSGESEDQSLTRFYACAGCARGPLPTDCCHQCAGCGPRKSDCLALHF 277  
 DB 199 CKGRSGESEDQSLTRFYACAGCARGPLPTDCCHQCAGCGPRKSDCLALHF 258  
 OY 278 NMSGICELHCPALVTYNTDFESMPNREGRTFGASCVTACPRYNYLSTDVGS 329  
 DB 259 NMSGICELHCPALVTYNTDFESMPNREGRTFGASCVTACPRYNYLSTDVGS 310

## RESULT 2

## FOHUGM

granulocyte-macrophage colony-stimulating factor precursor (validated) - human

N:Alternate names: colony-stimulating factor 2; GM-CSF

C:Species: Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 08-Dec-2000

C:Accession: C24636; 159065; A01853; A44175; JCI1090

R:Myakeake, S.; Otsuka, T.; Yokota, T.; Lee, F.; Aral, K.

EMBO J. 4, 2561-2568, 1985

A:Title: Structure of the chromosomal gene for granulocyte-macrophage colony stimulating

A:Reference number: A91015; M0ID:86030234; PMID:3876930

A:Accession: C24636

A:Molecule type: DNA

A:Residues: 1-144 <MY>

A:Cross-references: EMBL:X03021; NID:g31858; PIDN:CAA26822.1; PID:g31859

R:Kausanskyy, K.; O'Hara, P.J.; Berkner, K.; Segal, G.M.; Hagen, F.S.; Adamson, J.W.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 3101-3105, 1986  
 A:Title: Genomic cloning, characterization, and multilignage growth-promoting activity  
 C:Accession: 159065; M0ID:86205844; PMID:3486413  
 A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-144 <RES>  
 A:Cross-references: GB:M13207; NID:g181147; PIDN:AAA98768.1; PID:g181148  
 R:Cantrrell, M.A.; Anderson, D.; Cerretti, D.P.; Price, V.; McKereghan, K.; Tushinski,  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6250-6254, 1985  
 A:Title: Cloning, sequence, and expression of a human granulocyte/macrophage colony-s  
 A:Reference number: A25169; M0ID:85298329; PMID:3898082  
 A:Accession: A25169  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <CAN>  
 A:Cross-references: GB:M1734; NID:g181149; PIDN:AAA52122.1; PID:g181150  
 R:Lee, F.; Yokota, T.; Otsuka, T.; Gemmell, L.; Larson, N.; Luh, J.; Aral, K.; Rennic  
 Proc. Natl. Acad. Sci. U.S.A. 82, 4360-4364, 1985  
 A:Title: Isolation of cDNA for a human granulocyte-macrophage colony-stimulating fact  
 A:Reference number: A01853; M0ID:85242684; PMID:3925454  
 A:Accession: A01853  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <LEE>  
 A:Cross-references: GB:M11720; NID:g183363; PIDN:AAA52578.1; PID:g183364  
 R:Wong, G.G.; Wittek, J.S.; Temple, P.A.; Wilkens, K.M.; Leary, A.C.; Luxenberg, D.P.;  
 A:Clark, S.C.  
 Science 228, 810-815, 1985  
 A:Title: Human GM-CSF: molecular cloning of the complementary DNA and purification of  
 A:Reference number: A44175; M0ID:85218749; PMID:3923623  
 A:Accession: A44175  
 A:Molecule type: mRNA  
 A:Residues: 1-116, 'T', 118-144 <MON>  
 A:Cross-references: GB:M10663; NID:g181145; PIDN:AAA52121.1; PID:g181146  
 A:Note: parts of this sequence, including the amino end of the mature protein, were c  
 R:Wen, D.Y.; Huang, B.R.; Cal, L.W.; Si, J.Y.  
 Acta Biochim. Biophys. Sin. 25, 651-655, 1993  
 A:Title: Amplification of human granulocyte-macrophage colony-stimulating factor cDNA  
 A:Reference number: JCI1090  
 A:Accession: JCI1090  
 A:Molecule type: protein  
 A:Residues: 18-21, 'C', 23-96, 'L', 98-144 <WEN>  
 C:Genetics:  
 A:Gene: GDB:CSF2  
 A:Cross-references: GDB:119812; OMIM:138960  
 A:Map position: 5q23.2-5q31.1  
 A:Introns: 53/3; 67/3; 109/3  
 C:Function:  
 A:Description: stimulates the differentiation and proliferation of hematopoietic prog  
 C:Superfamily: granulocyte-macrophage colony-stimulating factor  
 C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-144/Product: granulocyte-macrophage colony-stimulating factor #status experiment  
 F:44,54/Binding site: carboxydrate (asn) (covalent) #status predicted

Query Match 18.4%; Score 127; DB 1; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 6,5e-118;  
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 549 APARSPSTQPMHVAIOEARLLNLSRDTAENMEVEISEMPDLOEPTCLOTRLE 608  
 DB 18 APARSPSTQPMHVAIOEARLLNLSRDTAENMEVEISEMPDLOEPTCLOTRLE 77  
 OY 609 LYKQGLSGSLTKLKGPLTMASHYKQCPPTPETSATQIITFESFENKDKFLVLPFD 668  
 DB 78 LYKQGLSGSLTKLKGPLTMASHYKQCPPTPETSATQIITFESFENKDKFLVLPFD 137  
 OY 669 CMEPVQD 675  
 DB 138 CMEPVQD 144

## RESULT 3

## I48161

P-185 precursor - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Jun-1999

C:Accession: I48161  
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagaoka, M.; Aral, M.; Yamazaki, Y.; Ishika

Gene 140, 251-255, 1994  
 A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.  
 A:Reference number: 148161; MUID:94193007; PMID:7908275  
 A:Accession: 148161  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1254 <RES>  
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595  
 C:Genetics:  
 A:Gene: neu  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif  
 Query Match 9.7%; Score 67; DB 2; Length 1254;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-57;  
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 250 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRT 309  
 |||  
 Db 231 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRT 290  
 |||  
 Oy 310 FGASCVT 316  
 |||  
 Db 291 FGASCVT 297  
 |||

RESULT 4  
 TVRTNU  
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 11-Jun-1999  
 C:Accession: A24562; A61204  
 R:Baragman, C.I.; Hung, M.C.; Weinberg, R.A.  
 Mature 319, 226-230, 1986  
 A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.  
 A:Reference number: A24562; MUID:86118662; PMID:3945311  
 A:Accession: A24562  
 A:Molecule type: mRNA  
 A:Residues: 1-1260 <EMBL>  
 A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746  
 R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen, C.  
 Carcinogenesis 12, 1975-1978, 1991  
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no p  
 2-thiazolyl]formamide or N-methyl-N-nitrosourea.  
 A:Reference number: A61204; MUID:92035293; PMID:1682063  
 A:Accession: A61204  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 637-663, 'V', 665-702 <MAS>  
 A:Note: authors translated the codon GCA for residue 25 as Val  
 C:Genetics:  
 A:Gene: neu  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>  
 F:658-680/Domain: transmembrane #status predicted <TMN>  
 F:723-988/Domain: protein kinase homology <KIN>  
 F:731-739/Region: protein kinase ATP-binding motif  
 F:71, 191, 263, 535, 576, 634/Binding site: carbonylrate (Asn) (covalent) #status predicted  
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted  
 F:758/Active site: lys #status predicted  
 F:882, 1227, 1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 7.5%; Score 52; DB 1; Length 1260;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-42;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 250 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESM 301  
 |||  
 Db 235 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESM 286  
 |||

RESULT 5  
 JH0610  
 acid phosphatase (EC 3.1.3.2) ACP precursor [validated] - human  
 N:Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphohydr  
 C:Species: Homo sapiens (man)  
 C:Date: 17-Aug-1992 #sequence\_revision 01-Dec-1995 #text\_change 08-Dec-2000  
 C:Accession: JH0610; J50693; A38608; S01331; A32419; S11147; S38863; S41251; S17042;  
 R:Shartel, F.S.; Li, S.S.L.  
 Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992  
 A:Title: Structure of human prostatic acid phosphatase gene.  
 A:Reference number: JH0610; MUID:92272747; PMID:1375464  
 A:Accession: JH0610  
 A:Molecule type: DNA  
 A:Residues: 1-386 <SHA>  
 A:Cross-references: GB:M97580; GB:M97581; GB:M97582; GB:M97583; GB:M97584; GB:M97585;  
 A:Accession: J50693  
 A:Molecule type: mRNA  
 A:Residues: 1-386 <SH3>  
 A:Cross-references: GB:M97589; NID:9189611; PIDN:AAA60021.1; PID:9189613  
 R:Van Etten, R.L.; Davidson, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.  
 J. Biol. Chem. 266, 2313-2319, 1991  
 A:Title: Covalent structure, disulfide bonding, and identification of reactive surfac  
 A:Reference number: A38608; MUID:91115848; PMID:1989985  
 A:Accession: A38608  
 A:Molecule type: mRNA  
 A:Residues: 1-386 <VAN>  
 A:Cross-references: GB:M24840; NID:9189620; PIDN:AAA6694.1; PID:9189621  
 A:Note: part of this sequence, including the amino end of the mature protein, was con  
 R:Vilko, P.; Virkkunen, P.; Henttu, P.; Roiko, K.; Solin, T.; Hultala, M.L.  
 FEBS Lett. 236, 275-281, 1988  
 A:Title: Molecular cloning and sequence analysis of cDNA encoding human prostatic aci  
 A:Reference number: S01331; MUID:86312881; PMID:2842184  
 A:Accession: S01331  
 A:Molecule type: mRNA  
 A:Residues: 1-14, 'A', 16, 'ASC', 20, 'CF', 23, 'C', 25-65, 'WIMPTHPA', 74-211, 'A', 213-386 <VIH  
 A:Cross-references: EMBL:X52174; NID:928321; PIDN:CAA36422.1; PID:928322  
 A:Note: part of this sequence, including the amino end of the mature protein, was con  
 R:Shartel, F.S.; Lee, H.; Leidenman, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.; Li, S  
 Biochem. Biophys. Res. Commun. 160, 79-86, 1989  
 A:Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein seq  
 A:Reference number: A32419; MUID:89228054; PMID:2712834  
 A:Accession: A32419  
 A:Molecule type: mRNA  
 A:Residues: 1-14, 'A', 16, 'ASC', 20, 'CF', 23, 'C', 25-94, 'D', 96-115, 'R', 117-214, 'S', 216-293  
 A:Cross-references: GB:M24902; NID:9189618; PIDN:AAA60022.1; PID:9189619  
 A:Note: the authors translated the codons GAC for residue 95 as Glu, CGT for residue  
 R:Tallor, P.G.; Govindan, M.V.; Patel, P.C.  
 Nucleic Acids Res. 18, 4928, 1990  
 A:Title: Nucleotide sequence of human prostatic acid phosphatase determined from a fu  
 A:Reference number: S11147; MUID:90370491; PMID:2395659  
 A:Accession: S11147  
 A:Status: preliminary; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-14, 'AFAC', 20, 'CF', 23, 'C', 25-45, 'H', 47-65, 'RIMPTHPA', 74-138, 'E', 140-156  
 A:Cross-references: EMBL:X53605; NID:935663; PIDN:CAA37673.1; PID:935684  
 R:Banas, B.; Blaschke, D.; Piltner, F.; Hoeriz, W.  
 submitted to the EMBL Data Library, April 1993  
 A:Description: Characterization of the promoter of the human prostatic acid phosphata  
 A:Reference number: S38863  
 A:Accession: S38863  
 A:Molecule type: DNA  
 A:Residues: 1-40 <BAN>  
 A:Cross-references: EMBL:X71391; NID:9416530; PIDN:CAA50514.1; PID:9416531  
 R:Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Blitt, E.; Porvari, K.; Taavitsalinen, P  
 submitted to the EMBL Data Library, September 1993  
 A:Description: Structural organization of human and rat prostate-specific acid phosph  
 sequence in the human gene promoter.  
 A:Reference number: S41251  
 A:Accession: S41251  
 A:Molecule type: DNA  
 A:Residues: 1-40 <VIR>

A:Cross-references: EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PID:g439658  
 R;Lee, H.; Chu, T.M.; Li, S.S.W.; Lee, C.  
 Biochem. J. 277, 759-765, 1991  
 A:Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.  
 A:Reference number: S17042; MUID:91336999; PMID:1908222  
 A:Accession: S17042  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 33-49 <LE>  
 R;Banas, B.; Blaschke, D.; Fittler, F.; Hoertz, W.  
 Biochim. Biophys. Acta 1217, 188-194, 1994  
 A:Title: Analysis of the promoter of the human prostatic acid phosphatase gene.  
 A:Reference number: S42730; MUID:94153995; PMID:8110833  
 A:Accession: S42730  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-40 <BA>  
 A:Cross-references: GB:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531  
 R;Sharief, F.S.; Li, S.S.  
 Biochem. Mol. Biol. Int. 33, 561-565, 1994  
 A:Title: Nucleotide sequence of human prostatic acid phosphatase ACP gene.  
 A:Reference number: I37175; MUID:95038536; PMID:7951074  
 A:Accession: I37175  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-386 <RS>  
 A:Cross-references: EMBL:U07097; NID:g515995; PIDN:AAB60640.1; PID:g515997  
 C:Comment: This protein is synthesized under androgen regulation by epithelial cells of C:Genetics:  
 A:Gene: GDB:ACPP  
 A:Cross-references: GDB:119644; OMIM:171790  
 A:Map position: 3q21.3-q25.2  
 A:Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2  
 C:Function:  
 A:Description: catalyzes the hydrolysis of a wide range of phosphate esters  
 C:Superfamily: mammalian acid phosphatase  
 C:Keywords: glycoprotein; phosphatidase; phosphoprotein; phosphoric monoester hydrolase  
 F:1-32/Domain: signal sequence #status predicted <SIG>  
 F:33-386/Product: acid phosphatase ACP #status experimental <MA>  
 F:44/86/Active site: Arg #status predicted  
 F:44/86/Active site: His (phosphatidase intermediate) #status predicted  
 F:94,220,333/Binding site: carboxylate (Asn) (covalent) #status experimental  
 F:161-372,215-313,347-351/Disulfide bonds: #status experimental

Query Match 5.18; Score 35; DB 1; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-26;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRAPLLARASLSIGFLFFFLFWRSLAKEL 35  
 Db 1 MRAPLLARASLSIGFLFFFLFWRSLAKEL 35

RESULT 6  
 A61632  
 granulocyte-macrophage colony-stimulating factor precursor - sheep  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 28-Jul-2000  
 C:Accession: A61632  
 R;O'Brien, P.M.; Rothel, J.S.; Sew, H.F.; Wood, P.R.  
 Immunol. Cell Biol. 69, 51-55, 1991  
 A:Title: Cloning and sequencing of the cDNA for ovine granulocyte-macrophage colony-stimulating factor  
 A:Reference number: A61632; MUID:91331592; PMID:1869289  
 A:Accession: A61632  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <OA>  
 A:Cross-references: GB:X55991; NID:g6983759; PIDN:CAA39463.1; PID:g6983760  
 C:Superfamily: granulocyte-macrophage colony-stimulating factor  
 F:1-17/Domain: signal sequence #status predicted <SIG>

Query Match 3.28; Score 22; DB 1; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 631 HYKHCPTPE 641  
 Db 100 HYKHCPTPE 110

RESULT 9  
 F0BOGM  
 granulocyte-macrophage colony-stimulating factor precursor - bovine  
 N:Alternate names: colony-stimulating factor 2; GM-CSF  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 27-Jan-1995  
 C:Accession: J10037

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 598 QEPICLQTRLELYKQGLRSLT 619  
 Db 67 QEPICLQTRLELYKQGLRSLT 88

RESULT 7  
 JH0469  
 granulocyte-macrophage colony-stimulating factor precursor - sheep  
 N:Alternate names: colony-stimulating factor 2; GM-CSF  
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
 C:Accession: JH0469; S16730  
 R;McInnes, C.J.; Haig, D.M.  
 Gene 105, 275-279, 1991  
 A:Title: Cloning and expression of a cDNA encoding ovine granulocyte-macrophage colony-stimulating factor  
 A:Reference number: JH0469; MUID:92039044; PMID:1937025  
 A:Accession: JH0469  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <MC1>  
 A:Cross-references: GB:X53561; NID:g1800; PIDN:CAA37632.1; PID:g1801  
 C:Comment: This protein is a glycoprotein cytokine produced and secreted by various c  
 C:Superfamily: granulocyte-macrophage colony-stimulating factor  
 C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-144/Product: granulocyte-macrophage colony-stimulating factor #status predicted  
 F:44/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 3.28; Score 22; DB 2; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 598 QEPICLQTRLELYKQGLRSLT 619  
 Db 67 QEPICLQTRLELYKQGLRSLT 88

RESULT 8  
 A44936  
 granulocyte-macrophage colony-stimulating factor precursor - dog  
 C:Species: Canis lupus familiaris (dog)  
 C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 16-Jul-1999  
 C:Accession: A44936  
 R;Nash, R.A.; Schuening, F.; Appelbaum, F.; Hammond, W.P.; Boone, T.; Morris, C.F.; S  
 Blood 78, 930-937, 1991  
 A:Title: Molecular cloning and in vivo evaluation of canine granulocyte-macrophage co  
 A:Reference number: A44936; MUID:91329842; PMID:1868252  
 A:Accession: A44936  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <NAS>  
 A:Cross-references: GB:S49738; NID:g233566; PIDN:AAB19466.1; PID:g233567  
 A:Note: sequence extracted from NCBI backbone (NCBI:49738, NCBI:P:49739)  
 C:Superfamily: granulocyte-macrophage colony-stimulating factor  
 F:1-17/Domain: signal sequence #status predicted <SIG>

Query Match 1.68; Score 11; DB 2; Length 144;  
 Best Local Similarity 100.0%; Pred. No. 0.0095;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 631 HYKHCPTPE 641  
 Db 100 HYKHCPTPE 110

R.Malliszewski, C.R.; Schoenborn, M.A.; Cerretti, D.P.; Wignall, J.M.; Picha, K.S.; Cosma  
Mol. Immunol. 25, 843-850, 1988  
A:Title: Bovine GM-CSF: molecular cloning and biological activity of the recombinant pro  
A:Reference number: J10037; MUID:89096971; PMID:3062386  
A:Accession: J10037

A:Status: preliminary  
A:Residues: 1-143 <MAL>  
C:Comment: This glycoprotein induces granulocyte, macrophage, and eosinophil colony form  
C:Superfamily: granulocyte-macrophage colony-stimulating factor  
C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell  
F:1-17/Domains: signal sequence #status predicted <SIG>  
F:18-143/Product: granulocyte-macrophage colony-stimulating factor #status predicted <MA  
F:44,54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.4% Score 10; DB 1; Length 143;  
Best Local Similarity 100.0%; Pred. No. 0.092;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 QEPCTQTRL 607  
DB 66 QEPCTQTRL 75

RESULT 10  
A42032  
epidermal growth factor receptor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Jun-1999  
C:Accession: A42032  
R:Flückinger, T.W.; Maibach, N.J.; Kung, H.J.  
Mol. Cell. Biol. 12, 883-893, 1992  
A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, t  
A:Reference number: A42032; MUID:92123214; PMID:1732751  
A:Accession: A42032

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-527 <FLI>  
A:Cross-references: GB:M77637; NID:9211737; PIDN:AAA48759.1; PID:9211738  
A:Experimental source: liver  
A:Note: sequence extracted from NCBI backbone (NCBI:76692, NCBI:76693)  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: ATP; growth factor receptor

Query Match 1.3% Score 9; DB 2; Length 527;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266  
DB 245 QCAAGCTGP 253

RESULT 11  
A36325  
epidermal growth factor receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 25-Jan-1991 #sequence\_revision 25-Jan-1991 #text\_change 10-Oct-1997  
C:Accession: A36325  
R:Perch, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Barr, H.S.  
Mol. Cell. Biol. 10, 2973-2982, 1990  
A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded b  
A:Reference number: A36325; MUID:90258888; PMID:2342466  
A:Accession: A36325  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-644 <PET>  
A:Cross-references: GB:M37394  
C:Superfamily: epidermal growth factor receptor; protein kinase homology  
C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 1.3% Score 9; DB 2; Length 644;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266  
DB 235 QCAAGCTGP 243

RESULT 12  
G0HUE  
epidermal growth factor receptor precursor - human  
N:Contents: protein-tyrosine kinase (EC 2.7.1.112) erbB  
C:Species: Homo sapiens (man)  
C:Date: 15-Nov-1984 #sequence\_revision 27-Nov-1985 #text\_change 11-Jun-1999  
C:Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05281; A60143;  
R:Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.  
rg, P.H.  
Nature 309, 418-425, 1984

A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression  
A:Reference number: A00641; MUID:84219729; PMID:6328312  
A:Accession: A00641

A:Molecule type: mRNA

A:Cross-references: EMBL:X00588; NID:931113; PIDN:CAA25240.1; PID:9757924  
A:Note: the authors translated the codon AAG for residue 540 as Asn  
R: Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985

A:Title: Characterization and sequence of the promoter region of the human epidermal  
A:Reference number: A25772; MUID:85270438; PMID:2991899  
A:Accession: A25772

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-29 <KSH>

A:Cross-references: GB:M11234; NID:9181981; PIDN:AAA52370.1; PID:9553272  
R: Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ullrich, A.; Waterfield, M.  
Oncogene Res. 1, 375-396, 1987

A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification  
A:Reference number: S30024; MUID:88217333; PMID:3329716  
A:Accession: S30024

A:Molecule type: DNA

A:Residues: 1-29 <HA2>  
A:Cross-references: EMBL:X06370; NID:931118; PIDN:CAA29668.1; PID:931119  
R: Haley, J.D.; Waterfield, M.D.  
J. Biol. Chem. 266, 1746-1753, 1991  
A:Title: Contributory effects of de Novo transcription and premature transcript termi  
A:Reference number: A38672; MUID:91107677; PMID:1988448  
A:Accession: A38672

A:Molecule type: DNA

A:Residues: 1-29 <HAL>

A:Cross-references: GB:M38425; NID:9181977; PIDN:AAA63171.1; PID:9553271  
A:Experimental source: carcinoma cell line A431-7  
R: Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;  
Nature 309, 806-810, 1984  
A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN  
A:Reference number: A00642; MUID:84245835; PMID:6330563  
A:Accession: A00642

A:Molecule type: mRNA

A:Residues: 150-187, 'NSVIOAV', 195, 'M', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-

A:Residues: 'RCAMRRA', 150-187, 'R', 813-942 <XUY>  
'798-799, 'ND', 802-811, 'R', 813-942 <XUY>  
A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF re  
R: Lin, C.R.; Chen, W.S.; Kruliger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma,  
Science 224, 843-848, 1984  
A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplificati  
A:Reference number: A43615; MUID:84196372; PMID:6326261  
A:Accession: A43615

A:Molecule type: mRNA

A:Residues: 713-964 <LIN>

A:Experimental source: epidermoid carcinoma cell line A431  
R: Stimmien, F.A.; Gope, M.L.; Schultz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
Biochem. Biophys. Res. Commun. 124, 125-132, 1984  
A:Reference number: A23062; MUID:85046483; PMID:6093780  
A:Accession: A23062

A:Molecule type: mRNA

A:Residues: 1028-1210 <SIM>  
R: Weber, W.; Gill, G.N.; Speiss, J.

Science 224, 294-297, 1984  
 A:Reference number: A05281; MUID:84172183; PMID:6324343  
 A:Accession: A05281  
 A:Molecule type: protein  
 R:Residues: 25-30,'S',32-51,45A-46T <MEB>  
 J. Biol. Chem. 260, 5205-5208, 1985  
 A:Title: Identification of residues in the nucleotide binding site of the epidermal growth factor receptor  
 A:Reference number: A60143; MUID:85182650; PMID:2985580  
 A:Accession: A60143  
 A:Molecule type: protein  
 A:Residues: 740-744,'X',746-747 <RUS>  
 R:Mroczkowski, B.; Mosig, G.; Cohen, S.  
 Nature 309, 270-273, 1984  
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superoxide dismutase  
 A:Reference number: A38023; MUID:84191554; PMID:6325948  
 A:Accession: A38023  
 A:Contents: annotation; receptor activity  
 A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA  
 R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.; Cell 59, 33-43, 1989  
 A:Title: Functional independence of the epidermal growth factor receptor from a domain I  
 A:Reference number: A33331; MUID:90003233; PMID:2750960  
 A:Accession: A33331  
 A:Contents: annotation; internalization signal  
 C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor complex  
 C:Genetics:  
 A:Gene: GDB:EGFR  
 A:Cross-references: GDB:120610; OMIM:131550  
 A:Map position: 7p12.3-7p12.1  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-1210/Product: EGF receptor #status predicted <EXT>  
 F:25-645/Domain: extracellular #status predicted <EXT>  
 F:75-300/Domain: EGF receptor extracellular domain repeat <EE1>  
 F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>  
 F:646-668/Domain: transmembrane #status predicted <TM>  
 F:669-1210/Domain: intracellular #status predicted <IM>  
 F:710-975/Domain: protein kinase homology <KIN>  
 F:718-726/Region: protein kinase ATP-binding motif  
 F:999-1046/Region: coated-pit mediated internalization signal  
 F:1047-1210/Region: inhibitory  
 F:128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:745/Active site: lys #status experimental  
 Query Match 1.38; Score 9; DB 1; Length 1210;  
 Best Local Similarity 100.0%; Pred. No. 6.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266  
 DB 235 QCAAGCTGP 243

RESULT 13  
 A53183  
 epidermal growth factor receptor precursor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999  
 C:Accession: A53183; A43818; S24942; A28941; S45325; I49643  
 R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.; Genes Dev. 8, 399-413, 1994  
 A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor tyrosine kinase  
 A:Reference number: A53183; MUID:94170986; PMID:8125255  
 A:Accession: A53183  
 A:Molecule type: mRNA  
 A:Residues: 1-1210 <UES>  
 A:Cross-references: GB:U03425  
 R:Aviv, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Gliori, D.; Morse, B.  
 Oncogene 6, 673-676, 1991  
 A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site  
 A:Reference number: A43818; MUID:91232866; PMID:2030916  
 A:Accession: A43818  
 A:Molecule type: mRNA

A:Residues: 1-714 <AVI>  
 A:Cross-references: GB:X59698  
 R:Eisinger, D.P.; Serrero, G.  
 submitted to the EMBL Data Library, June 1992  
 A:Reference number: S24942  
 A:Accession: S24942  
 A:Molecule type: mRNA  
 A:Residues: 969-971,'K',973-1115,'D' <EIS>  
 A:Cross-references: EMBL:Z12608  
 R:Heisermann, G.J.; Gill, G.N.  
 J. Biol. Chem. 263, 13152-13158, 1988  
 A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylate  
 A:Reference number: A28941; MUID:88330814; PMID:3138233  
 A:Accession: A28941  
 A:Molecule type: protein  
 A:Residues: 669-674,'X',696-704,'V',706-707,'989-992,'XX',995-996,'X',998-1000,1002-10  
 R:Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.  
 submitted to the EMBL Data Library, April 1994  
 A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor  
 A:Reference number: S45325  
 A:Accession: S45325  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-971,'K',973-1210 <VER>  
 A:Cross-references: EMBL:X78987; NID:9488830; PIDN:CAA55587.1; PID:9488831  
 R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Day, S.K.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993  
 A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse  
 A:Reference number: I49643; MUID:93126380; PMID:7678348  
 A:Accession: I49643  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 12-20,22-132 <RES>  
 A:Cross-references: GB:I06864; NID:9193001; PIDN:AAA53029.1; PID:9567201  
 C:Genetics:  
 A:Gene: EGFR  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phospho  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:648-670/Domain: transmembrane #status predicted <TM>  
 F:712-977/Domain: protein kinase homology <KIN>  
 F:720-728/Region: protein kinase ATP-binding motif  
 F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental  
 F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental  
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental  
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental  
 F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental  
 Query Match 1.38; Score 9; DB 2; Length 1210;  
 Best Local Similarity 100.0%; Pred. No. 6.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266  
 DB 235 QCAAGCTGP 243

RESULT 14  
 T4CHLV  
 epidermal growth factor receptor precursor - chicken  
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erpb  
 C:Species: Gallus gallus (chicken)  
 C:Date: 28-Feb-1986 #sequence\_revision 05-May-1995 #text\_change 04-Feb-2000  
 C:Accession: A27720; A00643  
 R:Lax, I.; Johnson, A.; Howk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Venn  
 Mol. Cell. Biol. 8, 1970-1978, 1988  
 A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in  
 A:Reference number: A27720; MUID:88261272; PMID:3260329  
 A:Accession: A27720  
 A:Molecule type: mRNA  
 A:Residues: 1-1223 <LAX>  
 A:Cross-references: GB:M20386  
 R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines

Cell 41, 719-726, 1985

A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro

A:Reference number: A00643; MUID:85228222; PMID:2988784

A:Accession: A00643

A:Molecule type: mRNA

A:Residues: 585-1223 <N1L>

A:Cross-references: GB:M10066

C:Genetics:

A:Gene: erbB

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-1223/Product: epidermal growth factor receptor #status predicted <EXT>

F:31-654/Domain: extracellular #status predicted <EXT>

F:61-307/Domain: EGF receptor extracellular domain repeat <EB1>

F:397-610/Domain: EGF receptor extracellular domain repeat <EB2>

F:655-677/Domain: transmembrane #status predicted <TM>

F:719-984/Domain: intracellular #status predicted <INT>

F:727-735/Region: protein kinase ATP-binding motif

F:136-202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #

F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted

F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F:754/Active site: Lys #status predicted

F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match

Best Local Similarity 1.3%; Score 9; DB 1; Length 1223;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTCP 266

DB 242 QCAAGCTCP 250

RESULT 15

I46269

granulocyte-macrophage colony stimulating factor - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 14-Feb-1997 #sequence-revision 14-Feb-1997 #text-change 16-Jul-1999

C:Accession: I46269

R:Smith, L.R.; Lundeen, K.A.; Diveley, J.P.; Carlo, D.J.; Brostoff, S.W.

Immunogenetics 39, 80, 1994

A:Title: Nucleotide sequence of the Lewis rat granulocyte-macrophage colony stimulating

A:Reference number: I46269; MUID:94041474; PMID:8225444

A:Accession: I46269

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-127 <SM1>

A:Cross-references: EMBL:U00620; NID:g392779; PIDN:AA18281.1; PID:g392780

C:Superfamily: granulocyte-macrophage colony-stimulating factor

Query Match

Best Local Similarity 1.2%; Score 8; DB 2; Length 127;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 609 LYKQGLRG 616

DB 61 LYKQGLRG 68

Search completed: April 28, 2003, 13:46:15  
Job time : 26 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:41:44 ; Search time 40 Seconds  
(without alignments)  
2298.573 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 690  
Sequence: 1 MRAPLLARASLSLGLF.....EPVQGAAPPPAAHHHHH 690

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: A\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	690	100.0	690	22	AAE13109 Human HER500-hgm-C
2	549	79.6	555	22	AAE13108 Human HER500-fusio
3	329	47.7	479	22	AAE13112 Human HER500-fgm-C
4	329	47.7	564	22	AAE13110 Human HER500-fusio
5	292	42.3	697	22	AAE13111 Human HER500-rgm-C
6	292	42.3	645	22	AAB60408 Human ErbB2 oncopr
7	292	42.3	645	22	AAB61593 Human ErbB2 extrac
8	292	42.3	653	21	AAB21200 Extracellular HER-
9	292	42.3	653	23	AA51145 Human Her-2/neu on
10	292	42.3	712	21	AAB31204 Human HER-2/neu fu

11	292	42.3	712	23	AA51149
12	292	42.3	782	18	AAW19764
13	292	42.3	919	21	AAB31203
14	292	42.3	919	23	AA51148
15	292	42.3	1200	21	AAB21208
16	292	42.3	1255	17	AAW01111
17	292	42.3	1255	20	AAW2406
18	292	42.3	1255	21	AAW21198
19	292	42.3	1255	21	AAW84780
20	292	42.3	1255	21	AAW2620
21	292	42.3	1255	22	AAE12130
22	292	42.3	1255	22	AAW85458
23	292	42.3	1255	22	AAW88267
24	292	42.3	1255	23	AAU77114
25	292	42.3	1255	23	AAU74545
26	292	42.3	1255	23	AAE20479
27	292	42.3	1255	23	AAW51143
28	292	42.3	1255	23	AAU77114
29	292	42.3	1255	23	AAU74545
30	289	41.9	289	22	AAE13120
31	289	41.9	951	21	AAW44993
32	277	40.1	1223	23	AAU98923
33	217	31.4	217	22	AAE13121
34	217	31.4	397	22	AAE13122
35	217	31.4	1179	22	AAE13123
36	173	25.1	191	20	AAW59354
37	173	25.1	191	20	AAW4572
38	173	25.1	191	22	AAB48767
39	166	24.1	166	19	AAW59345
40	166	24.1	166	22	AAB62074
41	166	24.1	166	22	AAB48763
42	158	22.9	1433	14	AAW39568
43	132	19.1	266	23	AAW11201
44	132	19.1	266	23	AAW51146
45	132	19.1	583	23	AAE20483

#### ALIGNMENTS

RESULT 1	
AAE13109	
ID	AAE13109 standard; Protein: 690 AA.
XX	
AC	AAE13109;
XX	
DT	28-JAN-2002 (first entry)
DE	
XX	
DE	Human HER500-hgm-CSF fusion protein construct.
XX	
KW	Immunostimulatory fusion protein; IFP: antigen component; therapy;
KW	immunostimulatory component; T-cell mediated immune response; DC;
KW	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW	PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW	membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW	HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW	HER500-hgm-CSF fusion protein.
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic.
XX	
PN	WO200174855-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US10515.
XX	
PR	30-MAR-2000; 2000US-193504P.
XX	
PA	(DEND-) DENDREON CORP.
XX	
PI	Laus R, Vidovic D, Gradis T;
XX	

Her-2/neu extracel  
Her-2/GM-CSF immuno  
Human HER-2/neu fu  
Her-2/neu extracel  
Human HER-2/neu pr  
Human HER-2/neu pr  
Human HER-2/neu on  
Human HER-2/neu pr  
Amino acid sequenc  
Human heregulin 2  
Human tyrosine kin  
Human HER-2/neu pr  
HER2/neu amino aci  
HER2 transgene pla  
Human Her-2 protei  
Human Her-2/neu pr  
Human Her-2/neu on  
Human Her-2/neu po  
Human HER2 (ErbB2)  
Mature human HER-2  
OC8scfv-erbB2EC fu  
Human breast cance  
Mature human HER-2  
NY-ESO-1C-HER-2 me  
Human SART3-IC-HER  
Human ErbB2 protei  
Human ErbB2 recept  
Human ErbB2 N-term  
Human ErbB2 domain  
ErbB2 protein doma  
Human ErbB2 extrac  
Sequence of c-erbB  
Human HER-2/neu pr  
Human Her-2/neu on  
Human protein. for

DR WPI: 2001-662965/76.  
 DR N-PSDB: AAD21565.  
 XX An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 26; 59pp: English.  
 XX  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC hbm-CSF fusion protein construct which comprises human PAP  
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2  
 CC intracellular domain, an Ala Ala linker, a mature human granulocyte-  
 CC macrophage colony stimulating factor (GM-CSF) sequence and a  
 CC C-terminal tag.  
 CC  
 SQ Sequence 690 AA:  
 Query Match 100.0%; Score 690; DB 22; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOYCTGDMKRLRPASPET 60  
 DB 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOYCTGDMKRLRPASPET 60  
 QY HLDMLRLHYGCGVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROYVFLQRLRI 120  
 DB HLDMLRLHYGCGVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROYVFLQRLRI 120  
 QY 61 HLDMLRLHYGCGVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROYVFLQRLRI 120  
 DB 61 HLDMLRLHYGCGVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROYVFLQRLRI 120  
 QY 121 VRGTQLEFEDNVYALVLDNGDPLNNTPTVGTASPGGLRELQRLSTELTKGGLVLIQRRNPOL 180  
 DB 121 VRGTQLEFEDNVYALVLDNGDPLNNTPTVGTASPGGLRELQRLSTELTKGGLVLIQRRNPOL 180  
 QY 181 CYODTILMKDIFHKNNDLALTLIDTNSRACHPCSPMKSGRCGSESDQSLTRIVCA 240  
 DB 181 CYODTILMKDIFHKNNDLALTLIDTNSRACHPCSPMKSGRCGSESDQSLTRIVCA 240  
 QY 181 CYODTILMKDIFHKNNDLALTLIDTNSRACHPCSPMKSGRCGSESDQSLTRIVCA 240  
 DB 181 CYODTILMKDIFHKNNDLALTLIDTNSRACHPCSPMKSGRCGSESDQSLTRIVCA 240  
 QY 241 GGCARCGPLPTDCHEQCAAGCTGPKHSDCLAEHFNHSGICELHCPALVTYNTDFEES 300  
 DB 241 GGCARCGPLPTDCHEQCAAGCTGPKHSDCLAEHFNHSGICELHCPALVTYNTDFEES 300  
 QY 301 MPNEGRYTFGASCVTACPYNYLSTDVSGAGCMVHHRHSSSTRSGGDLTLGLEPSEE 360  
 DB 301 MPNEGRYTFGASCVTACPYNYLSTDVSGAGCMVHHRHSSSTRSGGDLTLGLEPSEE 360  
 QY 361 EAPSPSLAPBEGAGSDVFDGDLGGAAGKGLQSLPTHRPSPLQRISEPTPLPSETGYV 420  
 DB 361 EAPSPSLAPBEGAGSDVFDGDLGGAAGKGLQSLPTHRPSPLQRISEPTPLPSETGYV 420  
 QY 421 APLTCSPOPEYVNPQVPPSPRPREPLPAARPAATLEAKTILSGKNGVYVDFVAFG 480  
 DB 421 APLTCSPOPEYVNPQVPPSPRPREPLPAARPAATLEAKTILSGKNGVYVDFVAFG 480  
 QY 481 GAVENPEYLTPOGGAPQHPHPAPFAFDNLXYWDDPPERGAPSTFGKPTAENPEY 540  
 DB 481 GAVENPEYLTPOGGAPQHPHPAPFAFDNLXYWDDPPERGAPSTFGKPTAENPEY 540  
 QY 541 LGIDVPAARARSPSPSTOWEHNNAIQEARRLNLSRQTAENNEVEYISEMFDQEP 600  
 DB 541 LGIDVPAARARSPSPSTOWEHNNAIQEARRLNLSRQTAENNEVEYISEMFDQEP 600  
 QY 601 TGLQTRLELTKGGLSLTKGFLTMASHYKOHCPPTETSCATQIITFESKENVKD 660  
 DB 601 TGLQTRLELTKGGLSLTKGFLTMASHYKOHCPPTETSCATQIITFESKENVKD 660

QY 661 FLTVIPDCMEYVQEGAPPPAANNHNNH 690  
 DB 661 FLTVIPDCMEYVQEGAPPPAANNHNNH 690  
 RESULT 2  
 ID AAE13108 standard; Protein: 555 AA.  
 AC AAE13108;  
 DT 28-JAN-2002 (first entry)  
 DE Human HER500 fusion protein construct.  
 KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW Immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human;  
 KW HER-2 protein; HER500 fusion protein.  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Synthetic.  
 PN WO200174855-A2.  
 PD 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US10515.  
 XX 30-MAR-2000; 2000US-193504P.  
 PA (DEND-) DENDREON CORP.  
 PI Laus R, Vidovic D, Graddis T;  
 DR WPI: 2001-662965/76.  
 DR N-PSDB: AAD21564.  
 PT An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 26; 59pp: English.  
 XX  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC fusion protein construct which comprises human PAP signal  
 CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal  
 CC sequence, mature HER-2 membrane distal extracellular and intracellular  
 CC domains and a C-terminal tag.  
 CC  
 SQ Sequence 555 AA:  
 Query Match 79.6%; Score 549; DB 22; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOYCTGDMKRLRPASPET 60  
 DB 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOYCTGDMKRLRPASPET 60  
 QY HLDMLRLHYGCGVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROYVFLQRLRI 120  
 DB HLDMLRLHYGCGVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROYVFLQRLRI 120

```

Oy 121 VRGTOLEFEDNVALAVLDNGDPLNNTTPVTGASPGGLRELOLSLREILKGVLIORNPOL 180
    |||||||
Db 121 VRGTOLEFEDNVALAVLDNGDPLNNTTPVTGASPGGLRELOLSLREILKGVLIORNPOL 180
Oy 181 CYODTILMKDIFRKNNOALATLIDNRSRACHPCSPMKGSKCMGESSDCCSLTRTYCA 240
    |||||||
Db 181 CYODTILMKDIFRKNNOALATLIDNRSRACHPCSPMKGSKCMGESSDCCSLTRTYCA 240
Oy 241 GGCACCKGRLPTDCCHEGCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTEFS 300
    |||||||
Db 241 GGCACCKGRLPTDCCHEGCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTEFS 300
Oy 301 MNPBGRTYFGASCYACPYNYLTSDVSGAGGMVHHRRSSSTSGGDLTLGLEPSEE 360
    |||||||
Db 301 MNPBGRTYFGASCYACPYNYLTSDVSGAGGMVHHRRSSSTSGGDLTLGLEPSEE 360
Oy 361 EAPRSPLAPSEGAGSDVEFDGLGMAKAGLOSLPTHDSPLQRYSEDPTVPLPSETDGYV 420
    |||||||
Db 361 EAPRSPLAPSEGAGSDVEFDGLGMAKAGLOSLPTHDSPLQRYSEDPTVPLPSETDGYV 420
Oy 421 APLTCSQPPEYVNOPDVRPQPSRREGPLPARPAGATLERAKTISPGKNGVYKVFARFAG 480
    |||||||
Db 421 APLTCSQPPEYVNOPDVRPQPSRREGPLPARPAGATLERAKTISPGKNGVYKVFARFAG 480
Oy 481 GAVENPEYLTPOGGAAPQPHPPAFSPAFDNLXYWDQDPREGAPPSTFKGPTAENPEY 540
    |||||||
Db 481 GAVENPEYLTPOGGAAPQPHPPAFSPAFDNLXYWDQDPREGAPPSTFKGPTAENPEY 540
Oy 541 LGLDVPAA 549
    |||||||
Db 541 LGLDVPAA 549

RESULT 3
AEI3112
ID AEI3112 standard; Protein: 479 AA.
XX
AC AEI3112;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER300-rGM-CSF fusion construct comprising OVA-derived peptide.
XX
KW Immunostimulatory fusion protein; IFP: antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW ovalbumin-derived octapeptide; OVA; rat; HER300-rGM-CSF fusion protein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Rattus norvegicus.
OS Chimeric - Unidentified.
XX
PN MO200174855-A2.
XX
DT 11-OCT-2001.
XX
DE 30-MAR-2001; 2001WO-US10515;
XX
PR 30-MAR-2000; 2000US-193504P.
XX
PA (DEND-) DENDREON CORP.
XX
PI Laus R, Vidovic D, Gradis T;
XX
DR WPI: 2001-662965/76.
XX
DR N-PSDB: AAD21568.
XX
PT An immunostimulatory fusion protein comprising the intracellular domain
of HER-2 and an antigen elicits an immune response to the antigen and

```

```

PT is useful for the treatment of associated cancer associated -
XX
PS Example 1; Page 27; 59pp; English.
XX
CC The invention relates to immunostimulatory fusion proteins (IFP) and
CC nucleic acid molecules encoding such proteins. The IFPs comprise a
CC polypeptide antigen component and an immunostimulatory component derived
CC from the intracellular domain of HER-2 protein which is effective to
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFP or superactivated dendritic cells
CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER300
CC rGM-CSF fusion protein construct which comprises human PAP
CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
CC signal sequence, mature HER-2 membrane distal extracellular domain,
CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
CC an Ala linker, a HER-2 membrane distal intracellular domain, a mature
CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence
CC and a C-terminal tag.
XX
SQ Sequence 479 AA:
Query Match 47.7%; Score 329; DB 22; Length 479;
Best Local Similarity 100.0%; Pred. No. 3,4e-293;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MRAAPLLARASLSLGFLEFFWLDRSVLAKELARGAASVOVCTGDMKRLPASPT 60
    |||||||
Db 1 MRAAPLLARASLSLGFLEFFWLDRSVLAKELARGAASVOVCTGDMKRLPASPT 60
Oy 61 HLDMLRHLYQGVQVQGNLELYLPTNASLSFLQDIOEVQGVYLLAHNOVQPLQRLRI 120
    |||||||
Db 61 HLDMLRHLYQGVQVQGNLELYLPTNASLSFLQDIOEVQGVYLLAHNOVQPLQRLRI 120
Oy 121 VRGTOLEFEDNVALAVLDNGDPLNNTTPVTGASPGGLRELOLSLREILKGVLIORNPOL 180
    |||||||
Db 121 VRGTOLEFEDNVALAVLDNGDPLNNTTPVTGASPGGLRELOLSLREILKGVLIORNPOL 180
Oy 181 CYODTILMKDIFRKNNOALATLIDNRSRACHPCSPMKGSKCMGESSDCCSLTRTYCA 240
    |||||||
Db 181 CYODTILMKDIFRKNNOALATLIDNRSRACHPCSPMKGSKCMGESSDCCSLTRTYCA 240
Oy 241 GGCACCKGRLPTDCCHEGCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTEFS 300
    |||||||
Db 241 GGCACCKGRLPTDCCHEGCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDTEFS 300
Oy 301 MNPBGRTYFGASCYACPYNYLTSDVGS 329
    |||||||
Db 301 MNPBGRTYFGASCYACPYNYLTSDVGS 329

RESULT 4
AEI3110
ID AEI3110 standard; Protein: 564 AA.
XX
AC AEI3110;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.
XX
KW Immunostimulatory fusion protein; IFP: antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; OVA;
KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
PN WO200174855-A2.

```

PD 11-OCT-2001.  
 XX 30-MAR-2001; 2001MO-US10515.  
 PF 30-MAR-2000; 2000US-193504P.  
 XX (DEND-) DENDREON CORP.  
 PA Laus R, Vidovic D, Graddis T;  
 PI WPI: 2001-662965/76.  
 DR N-PSDB; AAD21566.  
 XX  
 XX An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 26; 59pp: English.  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC fusion protein construct which comprises human PAP signal  
 CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal  
 CC sequence, mature HER-2 membrane distal extracellular domain, an  
 CC Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,  
 CC HER-2 membrane distal intracellular domain and a C-terminal tag.  
 XX  
 XX Sequence 564 AA:  
 SQ  
 Query Match 47.7%; Score 329; DB 22; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-293;  
 Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRAAPLLARAAASLSIGFLFLFWMFDRSVLAKELARGAASSTOYCTGTDKRLRPASPET 60  
 DB 1 MRAAPLLARAAASLSIGFLFLFWMFDRSVLAKELARGAASSTOYCTGTDKRLRPASPET 60  
 QY 61 HLDMLRLHYGCGVVOGNGLEETLPTNASLSFLDDIOEVGYVLIANOVROYPLQRLRI 120  
 DB 61 HLDMLRLHYGCGVVOGNGLEETLPTNASLSFLDDIOEVGYVLIANOVROYPLQRLRI 120  
 QY 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 DB 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 QY 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 DB 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 QY 181 CYODTILMKDIFHKNNQALTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLTRTVCA 240  
 DB 181 CYODTILMKDIFHKNNQALTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLTRTVCA 240  
 QY 241 GGCARCKGRLPTDCHEQCAAGCTGPKHSDCLACLFHNHSGICELHPALVYNTWTFES 300  
 DB 241 GGCARCKGRLPTDCHEQCAAGCTGPKHSDCLACLFHNHSGICELHPALVYNTWTFES 300  
 QY 301 MPNPEGRTYTGASCVTACPYNYLSTDVGS 329  
 DB 301 MPNPEGRTYTGASCVTACPYNYLSTDVGS 329  
 RESULT 5  
 AAE13111  
 ID AAE13111 standard; Protein; 697 AA.  
 XX  
 AC AAE13111;  
 XX  
 DT 28-JAN-2002 (first entry)  
 XX  
 DE Human HER500-rgm-CSF fusion construct comprising OVA-derived peptide.

XX  
 KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;  
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;  
 KW ovalbumin-derived octapeptide; OVA; rat; HER500-rgm-CSF fusion protein.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Rattus norvegicus.  
 OS Chimeric - Unidentified.  
 XX  
 PN WO200174855-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001MO-US10515.  
 XX  
 PR 30-MAR-2000; 2000US-193504P.  
 XX (DEND-) DENDREON CORP.  
 PA Laus R, Vidovic D, Graddis T;  
 PI WPI: 2001-662965/76.  
 DR N-PSDB; AAD21566.  
 XX  
 XX An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 27; 59pp: English.  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC rgm-CSF fusion protein construct which comprises human PAP  
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2  
 CC signal sequence, mature HER-2 membrane distal extracellular domain,  
 CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,  
 CC HER-2 membrane distal intracellular domain, an Ala Ala linker, a mature  
 CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence  
 CC and a C-terminal tag.  
 XX  
 XX Sequence 697 AA:  
 SQ  
 Query Match 47.7%; Score 329; DB 22; Length 697;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-293;  
 Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRAAPLLARAAASLSIGFLFLFWMFDRSVLAKELARGAASSTOYCTGTDKRLRPASPET 60  
 DB 1 MRAAPLLARAAASLSIGFLFLFWMFDRSVLAKELARGAASSTOYCTGTDKRLRPASPET 60  
 QY 61 HLDMLRLHYGCGVVOGNGLEETLPTNASLSFLDDIOEVGYVLIANOVROYPLQRLRI 120  
 DB 61 HLDMLRLHYGCGVVOGNGLEETLPTNASLSFLDDIOEVGYVLIANOVROYPLQRLRI 120  
 QY 61 HLDMLRLHYGCGVVOGNGLEETLPTNASLSFLDDIOEVGYVLIANOVROYPLQRLRI 120  
 DB 61 HLDMLRLHYGCGVVOGNGLEETLPTNASLSFLDDIOEVGYVLIANOVROYPLQRLRI 120  
 QY 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 DB 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 QY 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 DB 121 VRGTQLEFDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180  
 QY 181 CYODTILMKDIFHKNNQALTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLTRTVCA 240  
 DB 181 CYODTILMKDIFHKNNQALTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLTRTVCA 240  
 QY 241 GGCARCKGRLPTDCHEQCAAGCTGPKHSDCLACLFHNHSGICELHPALVYNTWTFES 300

Db 241 GGCACCKPPLPTDCCHEOCAGCTGPKHSDCLAHFNHSGICEHCPALVTYNTDTEFS 300  
OY 301 MPNPEGRTFGASCVTACPYNLTSDVGS 329  
Db 301 MPNPEGRTFGASCVTACPYNLTSDVGS 329

## RESULT 6

AAB60408 standard; Protein: 645 AA.

XX AAB60408;

DT 24-APR-2001 (first entry)

XX Human ErbB2 oncoprotein, SEQ ID NO:13.

XX Anti-ErbB2 monoclonal antibody 2C4; HER2; mouse; murine; humanised; VL;  
XX light chain variable region; cancer; cytostatic; EGFR-expressing cancer;  
XX epidermal growth factor receptor; colon cancer; rectal cancer; tumour;  
XX colorectal cancer; non-small cell lung cancer; metastatic breast cancer;  
XX affinity purification.

XX Homo sapiens.

XX WO200100245-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-US17366.

XX 25-JUN-1999; 99US-0141316.

XX (GETH ) GENENTECH INC.

XX Adams CW, Presta LG, Sliwkowsky M;

XX WPI; 2001-080862/09.

PT Treating cancer in a human, where the cancer expresses epidermal growth  
factor receptor (EGFR), comprises administering an antibody which binds  
ErbB2 -

XX Example 1; Fig 1A; 89pp; English.

XX The invention relates to a method for treating cancer in a human patient,  
XX wherein the cancer expresses epidermal growth factor receptor (EGFR),  
XX comprising administering an antibody which binds ErbB2 (HER2; AAB60408).  
XX In particular, the anti-ErbB2 antibody is the murine monoclonal antibody  
XX 2C4 (AAB60396, AAB60397) or a humanised version of 2C4 (AAB60398,  
XX AAB60399). The invention also encompasses an isolated nucleic acid  
XX encoding a humanised ErbB2-binding antibody; vectors and host cells  
XX comprising such nucleic acids; the recombinant production of a humanised  
XX ErbB2-binding antibody; and an immunocjugate comprising a humanised  
XX ErbB2-binding antibody and a cytotoxic drug. The ErbB2-binding antibodies  
XX act by antagonising ErbB receptors, and as inhibitors of transforming  
XX growth factor alpha (TGF-alpha)-activated mitogen activated protein  
XX kinase (MAPK). The method of the invention is used for treating cancer,  
XX especially colon cancer, rectal cancer, colorectal cancer, lung cancer  
XX (especially non-small cell lung cancer), or breast cancer (especially  
XX metastatic breast cancer). The antibodies may also have non-therapeutic  
XX uses e.g., as affinity purification agents. Using an antibody which binds  
XX to ErbB2 to treat cancer is preferable to the use of EGFR-targeted  
XX drugs, as EGFR is also highly expressed in other tissues such as the  
XX liver and skin, where the active drug will also bind, with skin toxicity  
XX having been observed for EGFR-targeted drugs. Antibodies which bind  
XX ErbB2 are anticipated to have a better safety profile than such drugs.  
XX The present sequence represents human ErbB2.

XX Sequence 645 AA;

Query Match 42.3%, Score 292; DB 22; Length 645;

Best Local similarity 100.0%; Pred. No. 4e-259;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAASQVCTGDMKRLPASPETHIDMLRHLYCGGVVGNLELYLPTNASTGLDIO 97  
Db 19 GAASQVCTGDMKRLPASPETHIDMLRHLYCGGVVGNLELYLPTNASTGLDIO 78  
OY 98 EVQGYVLLAHNQVROPPLRLRYVGTQLFEDNYALAVLDNGDPLNNTPTVTSRGLR 157  
Db 79 EVQGYVLLAHNQVROPPLRLRYVGTQLFEDNYALAVLDNGDPLNNTPTVTSRGLR 138  
OY 158 ELQRLSTLEILKGVLIQPNPOLCYQDTLLMKDIFHKNNQALATLIDITNRSRACPCSPM 217  
Db 139 ELQRLSTLEILKGVLIQPNPOLCYQDTLLMKDIFHKNNQALATLIDITNRSRACPCSPM 198  
OY 218 CKGSRCKNCESSDQSLRTVTCAGGCARCKGLPTDCCHEOCAGCTGPKHSDCLAHF 277  
Db 199 CKGSRCKNCESSDQSLRTVTCAGGCARCKGLPTDCCHEOCAGCTGPKHSDCLAHF 258  
OY 278 NMSGICEHCPALVTYNTDTEFSMPNPGRTFGASCVTACPYNLTSDVGS 329  
Db 259 NMSGICEHCPALVTYNTDTEFSMPNPGRTFGASCVTACPYNLTSDVGS 310

## RESULT 7

AAB61593 standard; protein: 645 AA.

XX AAB61593;

DT 04-APR-2001 (first entry)

XX Human ErbB2 extracellular domain.

XX Human; ErbB2; cytostatic; prostate cancer; receptor tyrosine kinase;  
XX antibody; ErbB receptor; monoclonal antibody 2C4; variable light chain.

XX Homo sapiens.

XX WO200100238-A1.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-US17423.

XX 25-JUN-1999; 99US-0141315.

XX (GETH ) GENENTECH INC.

XX (SLOK ) SLOAN KETTERING INST CANCER RES.

XX Agus DB, Scher HI, Sliwkowski MX;

XX WPI; 2001-159131/16.

PT Treating prostate cancer in a human comprises administering an antibody  
which binds ErbB2 and blocks ligand activation of an ErbB receptor -

XX Disclosure; Fig 1; 93pp; English.

XX The ErbB family of receptor tyrosine kinases are important mediators of  
XX cell growth, differentiation and survival. The receptor family includes  
XX four distinct members including Epidermal Growth Factor Receptor (EGFR or  
XX ErbB1), HER2 (ErbB2 or p185<sup>neu</sup>), HER3 (ErbB3) and Her4 (ErbB4 or tyro2).  
XX The present invention relates to a method for treating prostate cancer.  
XX The method comprises administering an antibody which binds ErbB2 and  
XX blocks ligand activation of an ErbB receptor. Preferably, the antibody  
XX blocks binding of monoclonal antibody 2C4 to ErbB2 and/or blocks  
XX TGF-alpha activation of mitogen-activated protein kinase (MAPK). The  
XX present sequence is the extracellular domain of human ErbB2.

XX Sequence 645 AA;

Query Match 42.3%, Score 292; DB 22; Length 645;

	Best Local Similarity	100.0%	Pred. No. 4e-259;	
Matches	292;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0
OY	38	GAASTOVCTGTDMLRLPASPTHTDMLRHLYOGCCVVOGNELELYLPTNASTSLFDIO	97	
Db	19	GAASTOVCTGTDMLRLPASPTHTDMLRHLYOGCCVVOGNELELYLPTNASTSLFDIO	78	
OY	98	EVQGVVLLAHNOVROPRLRIIVGTOLEEDNYALAVLNDNDPLNNTTPVVGASGGIR	157	
Db	79	EVQGVVLLAHNOVROPRLRIIVGTOLEEDNYALAVLNDNDPLNNTTPVVGASGGIR	138	
OY	158	ELQSLSTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQLATLTLDTNRSRACHPCSPM	217	
Db	139	ELQSLSTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQLATLTLDTNRSRACHPCSPM	198	
OY	218	CKGSRCKWGSSESDCOSLRTTVCAGGCARCKGPLPTDCCHEQCAAGCTGKHSDCIACLHF	277	
Db	199	CKGSRCKWGSSESDCOSLRTTVCAGGCARCKGPLPTDCCHEQCAAGCTGKHSDCIACLHF	258	
OY	278	NHSGICELCPALVLYYNTDTFESMPNDEGRYRFGASCYACAPYNYLSTVGS	329	
Db	259	NHSGICELCPALVLYYNTDTFESMPNDEGRYRFGASCYACAPYNYLSTVGS	310	

RESULT 8  
AAB21200

ID AAB21200 standard; protein; 653 AA.

AC AAB21200;

DT 12-JAN-2001 (first entry)

DE Extracellular HER-2/neu protein.

HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine; ...

KW colon cancer.

OS Unidentified.

PN WO200044899-

PD 03-AUG-2000.

PF 28-JAN-2000; 2000WO-US02164.

PR 29-JAN-1999; 99US-0117976.

PA (CORI-) CORIXA CORP.

2

**XX**

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

PT useful for vaccinating against breast, ovarian, colon, lung and

2  
1  
0  
1  
0  
3  
3  
4  
1  
1

CC The present sequence is the extracellular HER-2/neu protein. HER-2/neu is  
CC a member of the tyrosine kinase family of receptor-like glycoproteins and  
CC shows homology to the epidermal growth factor receptor (EGFR). It  
CC probably plays a part in cell growth and/or differentiation. The  
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising a  
CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation  
CC domain may be used to treat or prevent cancer by eliciting or  
CC enhancing an immune response to the HER-2/neu protein. It may be used  
CC to treat malignancies such as breast, ovarian, colon, lung and  
CC prostate cancers, and may be used as an antigen to vaccinate against  
CC these neoplasias.

50 Sequence 653 AA;

Query Match	42.38;	Score 292;	DB 21;	Length 653;
PostgreSQL Statement	100.00;	Exec No	4	10-350

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 9

ID	AA
AAM51145	standard; protein; 653 AA

AC AAM51145;

DT 17-JUN-2002 (first entry)

DE . Human Her-2/neu oncoprotein extracellular domain

KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;

XX  
Hence

XX  
DN 100300313341-83

XX 14-FEB-2003  
DD

XX 03-DEC-2001 2001FQ-FRC34293

XX 03-AUG-2000 2000UG-0632E02

XX  
XX  
COPT- COPTA COPEPA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS  
 VV

PI Cheever MA, Gheysen D, VY

DR WPI; 2002-241743/29  
xx

Her-2/neu fusion protein for treating or preventing cancer by eliciting

PT extracellular domain fused to Her-2/neu intracellular or

[illegible]

CC The present sequence is that of the extracellular domain of  
CC human Her-2/neu (p185 glycoprotein or c-erbB2), an oncogenic  
CC self-protein and target for anti-cancer vaccines. The Her-2/neu  
CC gene is amplified and p185 is overexpressed in a variety of cancers  
CC including breast, ovarian, colon, lung and prostate cancer.  
CC Her-2/neu (see AA551143) is a member of the tyrosine kinase family  
CC of receptor-like glycoproteins. It comprises an extracellular  
CC domain with homology to the epidermal growth factor receptor  
CC (EGFR), a highly hydrophobic transmembrane domain and a C-terminal

CC Intracellular domain that also shows homology to EGFR. Its  
CC overexpression correlates with a poor prognosis in breast and  
CC ovarian cancers. The invention provides Her-2/neu fusion  
CC proteins, nucleic acids encoding them, viral vectors, and vaccines  
CC comprising the fusion proteins or nucleic acid molecules. In  
CC preferred fusion proteins, the extracellular domain of a Her-2/neu  
CC protein is fused to a Her-2/neu intracellular domain or  
CC phosphorylation domain (or its delta-p fragment). An immune  
CC response to Her-2/neu protein is elicited or enhanced by  
CC administering the fusion protein in the form of a vaccine, or by  
CC transfecting cells of an animal *ex vivo* with a nucleic acid  
CC encoding the fusion protein, and delivering the transfected cells  
CC to the animal. The fusion proteins, nucleic acids, and isolated  
CC specific T-cells are useful for inhibiting the development of a  
CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
CC in a patient. T cells that specifically react with a Her-2/neu  
CC fusion protein can be used to remove tumour cells from a sample in  
CC order to inhibit the development of cancer in a patient.

XX Sequence 653 AA;

Query Match 42.3%; Score 292; DB 23; Length 653;

Best Local Similarity 100.0%; Pred. No. 4,1e-259; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGDMKRLPASPETHLDMRLHLYGCGVVGNNLELYLPTNLSLFLDIO 97  
DB 19 GAATGCTGDMKRLPASPETHLDMRLHLYGCGVVGNNLELYLPTNLSLFLDIO 78  
QY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDGDDPLNNTPTVTGASPGGLR 157  
DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDGDDPLNNTPTVTGASPGGLR 138  
QY 158 ELQRLSLLEILKGGVLIQNRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217  
DB 139 ELQRLSLLEILKGGVLIQNRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198  
QY 218 CKGRKMGESSEDCOSLTRIVCAGGACARCKGRLPTDCHECCAGACTGPKHSDCIACLHF 277  
DB 199 CKGRKMGESSEDCOSLTRIVCAGGACARCKGRLPTDCHECCAGACTGPKHSDCIACLHF 258  
QY 278 NHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329  
DB 259 NHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 10

AAB21204 AAB21204 standard; protein; 712 AA.

XX AAB21204;

XX 12-JAN-2001 (first entry)

XX Human HER-2/neu fusion protein.

XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

KW breast cancer; prostate cancer; ovarian cancer; lung cancer;

KW colon cancer; fusion protein.

XX Homo sapiens.

XX Synthetic.

XX WO200044899-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US02164.

XX 29-JAN-1999; 99US-0117976.

XX (CORI-) CORIYA CORP.

XX (SMIK) SMITHKLINE BEECHAM.

XX Cheever MA, Gheysen D;

XX WPI: 2000-505976/45.

XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins

XX useful for vaccinating against breast, ovarian, colon, lung and

XX prostate cancers -

XX Claim 27; Fig 13; 128pp; English.

The present sequence is a fusion protein comprising the extracellular  
domain and a preferred portion of the phosphorylation domain of the human  
HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase family of  
receptor-like glycoproteins and shows homology to the epidermal growth  
factor receptor (EGFR). It probably plays a part in cell growth and/or  
differentiation. The HER-2/neu gene is an oncogene. HER-2/neu fusion  
proteins may be used to treat or prevent cancer by eliciting or enhancing  
an immune response to the HER-2/neu protein. They may be used to treat  
malignancies such as breast, ovarian, colon, lung and prostate cancers,  
and may be used as an antigen to vaccinate against these neoplasias.

XX Sequence 712 AA;

Query Match 42.3%; Score 292; DB 21; Length 712;

Best Local Similarity 100.0%; Pred. No. 4,4e-259; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGDMKRLPASPETHLDMRLHLYGCGVVGNNLELYLPTNLSLFLDIO 97  
DB 19 GAATGCTGDMKRLPASPETHLDMRLHLYGCGVVGNNLELYLPTNLSLFLDIO 78  
QY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDGDDPLNNTPTVTGASPGGLR 157  
DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNVALAVLDGDDPLNNTPTVTGASPGGLR 138  
QY 158 ELQRLSLLEILKGGVLIQNRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217  
DB 139 ELQRLSLLEILKGGVLIQNRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198  
QY 218 CKGRKMGESSEDCOSLTRIVCAGGACARCKGRLPTDCHECCAGACTGPKHSDCIACLHF 277  
DB 199 CKGRKMGESSEDCOSLTRIVCAGGACARCKGRLPTDCHECCAGACTGPKHSDCIACLHF 258  
QY 278 NHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329  
DB 259 NHSGICELHCPALVYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 11

AAM51149 AAM51149 standard; protein; 712 AA.

XX AAM51149;

XX 17-JUN-2002 (first entry)

XX Her-2/neu extracellular domain-delta-phosphorylation domain fusion.

XX Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;

KW tyrosine kinase; receptor; c-erbB2; gene therapy.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

XX Domain

XX Domain

XX WO200212341-A2.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-US24283.  
 PF 03-AUG-2000; 2000US-0632507.  
 XX  
 PR  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Cheever MA, Gheysen D;  
 XX  
 DR WPI; 2002-241743/29.  
 XX  
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting  
 PT or enhancing an immune response to the protein, has Her-2/neu  
 PT extracellular domain fused to Her-2/neu intracellular or  
 PT phosphorylation domain  
 XX  
 PS Claim 37; Fig 13; 141pp; English.  
 XX  
 CC The present sequence is that of a fusion protein between the  
 CC extracellular domain and a fragment (DeltapD) of the phosphorylation  
 CC domain of human Her-2/neu (see AAM51143), an oncogenic self-protein  
 CC and target for anti-cancer vaccines. The fusion protein can be  
 CC obtained by recombinant DNA methods. Her-2/neu overexpression  
 CC correlates with a poor prognosis in breast and ovarian cancers.  
 CC The invention provides Her-2/neu fusion proteins, nucleic acids  
 CC encoding them, viral vectors, and vaccines comprising the fusion  
 CC proteins or nucleic acid molecules. In preferred fusion proteins,  
 CC the extracellular domain of Her-2/neu is fused to a Her-2/neu  
 CC intracellular domain or phosphorylation domain (or its DeltapD  
 CC fragment). An immune response to Her-2/neu protein is elicited or  
 CC enhanced by administering the fusion protein in the form of a vaccine,  
 CC or by transfecting cells of an animal *ex vivo* with a nucleic acid  
 CC encoding the fusion protein, and delivering the transfected cells  
 CC to the animal. The fusion proteins, nucleic acids, and isolated  
 CC specific T-cells are useful for inhibiting the development of a  
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer  
 CC in a patient. T cells that specifically react with a Her-2/neu  
 CC fusion protein can be used to remove tumour cells from a sample in  
 CC order to inhibit the development of cancer in a patient.  
 CC  
 XX  
 SQ Sequence 712 AA;  
 Query Match 42.3%; Score 292; DB 23; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-259;  
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 38 GAASTOYCTGDMKRLRPASPETHLDMRLHYGCGVVGNGLETTYLPTNASLSFLDIO 97  
 DB 19 GAASTOYCTGDMKRLRPASPETHLDMRLHYGCGVVGNGLETTYLPTNASLSFLDIO 78  
 QY 98 EVGGYVLIANNOVROVPLQRLRIYRGTLFEDNYALAVLNDGDP LNNTTPTGASPGGLR 157  
 DB 79 EVGGYVLIANNOVROVPLQRLRIYRGTLFEDNYALAVLNDGDP LNNTTPTGASPGGLR 138  
 QY 158 ELQIRSLTEILKGVLIQIRNPOLCYODTIIMKDI FHKNNOLATLIDTNSRACHPCSPM 217  
 DB 139 ELQIRSLTEILKGVLIQIRNPOLCYODTIIMKDI FHKNNOLATLIDTNSRACHPCSPM 198  
 QY 218 CKGSRGCGESSEDCOSLTRIVCAGGCAKCGPLPTDCCHQCAAGCTGPHNSDCLACHF 277  
 DB 199 CKGSRGCGESSEDCOSLTRIVCAGGCAKCGPLPTDCCHQCAAGCTGPHNSDCLACHF 258  
 QY 278 NHSICELHCPALVTYNTDFESMPNPEGRTTFCASCVTACPYNYLSTDVGS 329  
 DB 259 NHSICELHCPALVTYNTDFESMPNPEGRTTFCASCVTACPYNYLSTDVGS 310  
 RESULT 12  
 AAM19764  
 ID AAM19764 standard; Protein; 782 AA.  
 XX  
 AC AAM19764;

XX 17-SEP-1997 (first entry)  
 DT  
 XX  
 DE Her2-GM-CSF immunostimulant fusion protein.  
 XX  
 KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;  
 KW growth factor receptor; oncogene; immunostimulant; cancer;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..653  
 FT Peptide /label= Her2  
 FT /label= 654..655  
 FT /label= Linker  
 FT Protein 656..782  
 FT /label= GM-CSF  
 XX  
 PN WO9724438-A1.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PF 23-DEC-1996; 96WO-US20241.  
 XX  
 PR 28-DEC-1995; 95US-0579823.  
 XX  
 PA (ACTI-) ACTIVATED CELL THERAPY INC.  
 XX  
 PI Laus R, Ruegg CL, Wu H;  
 XX  
 DR WPI; 1997-363674/33.  
 DR N-PSDB; AAT72725.  
 XX  
 PT Potent APC that activates T-cells to give multivalent cellular  
 PT immune response - can also induce a cytotoxic T-cell response in a  
 PT vertebrate subject  
 XX  
 PS Disclosure; Fig 8; 45pp; English.  
 XX  
 CC A fusion protein (AAM19764) comprises Her2 (a growth factor receptor  
 CC that is overexpressed in breast, ovarian and other cancer cells)  
 CC and granulocyte-macrophage colony stimulating factor (GM-CSF). It  
 CC is the expression product of a nucleic acid molecule (AAT72725)  
 CC prep. by PCR amplification of Her2 cDNA from a breast cancer cell  
 CC line and fusion to GM-CSF cDNA. Fusion expression vectors can be  
 CC used to transfect mammalian and insect cells. The Her2-GM-CSF  
 CC fusion protein is used to generate anti-Her2 immunity. Tumour  
 CC cells are eliminated by cytotoxic T lymphocytes activated *in vivo*  
 CC or *in vitro* by exposure to antigen-presenting cells exposed to the  
 CC fusion protein.  
 CC  
 XX  
 SQ Sequence 782 AA;  
 Query Match 42.3%; Score 292; DB 18; Length 782;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-259;  
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 38 GAASTOYCTGDMKRLRPASPETHLDMRLHYGCGVVGNGLETTYLPTNASLSFLDIO 97  
 DB 19 GAASTOYCTGDMKRLRPASPETHLDMRLHYGCGVVGNGLETTYLPTNASLSFLDIO 78  
 QY 98 EVGGYVLIANNOVROVPLQRLRIYRGTLFEDNYALAVLNDGDP LNNTTPTGASPGGLR 157  
 DB 79 EVGGYVLIANNOVROVPLQRLRIYRGTLFEDNYALAVLNDGDP LNNTTPTGASPGGLR 138  
 QY 158 ELQIRSLTEILKGVLIQIRNPOLCYODTIIMKDI FHKNNOLATLIDTNSRACHPCSPM 217  
 DB 139 ELQIRSLTEILKGVLIQIRNPOLCYODTIIMKDI FHKNNOLATLIDTNSRACHPCSPM 198  
 QY 218 CKGSRGCGESSEDCOSLTRIVCAGGCAKCGPLPTDCCHQCAAGCTGPHNSDCLACHF 277  
 DB 199 CKGSRGCGESSEDCOSLTRIVCAGGCAKCGPLPTDCCHQCAAGCTGPHNSDCLACHF 258





XX Sequence 919 AA;  
SQ Query Match 42.3%; Score 292; DB 23; Length 919;  
Best Local Similarity 100.0%; Pred. No. 5.4e-259;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGCTDMKRLRPASPETHLDMRLHYGCGVVGNNLELYLPTNASLSFLDIO 97  
DB 19 GAATGCTGCTDMKRLRPASPETHLDMRLHYGCGVVGNNLELYLPTNASLSFLDIO 78  
QY 98 EVGGVLLAHNQVROVPLQRLRIYRGTOLEFDNYALAVLDGDDPLNNTPTVTGASPGGLR 157  
DB 79 EVGGVLLAHNQVROVPLQRLRIYRGTOLEFDNYALAVLDGDDPLNNTPTVTGASPGGLR 138  
QY 158 ELQRLSLTEILKGGVLIQIRNPOLCYODTIILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217  
DB 139 ELQRLSLTEILKGGVLIQIRNPOLCYODTIILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198  
QY 218 CKGRMCWESSEDCOSLTRIVCAGGACARCKGPLPTDCHEOCAGCTGPKHSDCLACLHF 277  
DB 199 CKGRMCWESSEDCOSLTRIVCAGGACARCKGPLPTDCHEOCAGCTGPKHSDCLACLHF 258  
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 329  
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 310

RESULT 15  
AAB21208  
ID AAB21208 standard; Protein: 1200 AA.  
XX AAB21208;  
AC 12-JAN-2001 (first entry)  
XX 12-JAN-2001 (first entry)  
XX Human HER-2/neu protein.  
DE Human HER-2/neu protein.  
XX Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;  
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
KW colon cancer.  
XX Homo sapiens.  
OS Homo sapiens.  
XX MO200044899-A1.  
PN 03-AUG-2000.  
PD 28-JAN-2000; 2000MO-US02164.  
XX 28-JAN-2000; 2000MO-US02164.  
PE 29-JAN-1999; 99US-0117976.  
XX 29-JAN-1999; 99US-0117976.  
PA (CORI-) CORIXA CORP.  
PA (SMIK) SMITHKLINE BEECHAM.  
XX Cheever MA, Gheysen D;  
PI WPI: 2000-505976/45.  
XX N-PSDB: AAA89736.  
DR N-PSDB: AAA89736.  
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
PT useful for vaccinating against breast, ovarian, colon, lung and  
PT prostate cancers -  
PS Disclosure: Fig 15; 128bp; English.  
XX The present sequence is the human HER-2/neu protein. It is a member  
CC of the tyrosine kinase family of receptor-like glycoproteins and shows  
CC homology to the epidermal growth factor receptor (EGFR). It probably  
CC plays a part in cell growth and/or differentiation. The HER-2/neu  
CC gene is an oncogene. An HER-2/neu fusion protein comprising a  
CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation  
CC domain may be used to treat or prevent cancer by eliciting or

CC enhancing an immune response to the HER-2/neu protein. It may be used  
CC to treat malignancies such as breast, ovarian, colon, lung and  
CC prostate cancers, and may be used as an antigen to vaccinate against  
CC these neoplasias.  
XX  
SQ Sequence 1200 AA;  
Query Match 42.3%; Score 292; DB 21; Length 1200;  
Best Local Similarity 100.0%; Pred. No. 6.8e-259;  
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAATGCTGCTDMKRLRPASPETHLDMRLHYGCGVVGNNLELYLPTNASLSFLDIO 97  
DB 19 GAATGCTGCTDMKRLRPASPETHLDMRLHYGCGVVGNNLELYLPTNASLSFLDIO 78  
QY 98 EVGGVLLAHNQVROVPLQRLRIYRGTOLEFDNYALAVLDGDDPLNNTPTVTGASPGGLR 157  
DB 79 EVGGVLLAHNQVROVPLQRLRIYRGTOLEFDNYALAVLDGDDPLNNTPTVTGASPGGLR 138  
QY 158 ELQRLSLTEILKGGVLIQIRNPOLCYODTIILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217  
DB 139 ELQRLSLTEILKGGVLIQIRNPOLCYODTIILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198  
QY 218 CKGRMCWESSEDCOSLTRIVCAGGACARCKGPLPTDCHEOCAGCTGPKHSDCLACLHF 277  
DB 199 CKGRMCWESSEDCOSLTRIVCAGGACARCKGPLPTDCHEOCAGCTGPKHSDCLACLHF 258  
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 329  
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 310

Search completed: April 28, 2003, 13:44:37  
Job time : 43 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 8, 2003, 12:36:28 ; Search time 210 Seconds  
(Without alignments)  
4083.173 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739  
Sequence: 1 MRAPILLARASLSIGLE.....EPYQEGAPRRPAHHNNHH 690

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL-frame.p2n.model -DEV-xlp  
-Q/cg2\_1/USPTO.spool/US09821883/runatc\_28042003\_104750\_14608/app\_query.fasta.1.839  
-DB-published.Applications.NA -QFMT-fastap -SUFIX-p2n.rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62  
-TRANS-human40.ccd -LIST=45 -DOCCALIGN=200 -THR SCORE-pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT-pio -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MALIGN=200000000 -USER=US09821883.ecgn\_1.1.80.etrnatc\_28042003\_104750\_14608  
-NCU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOS  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Published Applications\_NA:\*  
1: /cg2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cg2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/ptodata/1/pubpna/PCUS\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cg2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cg2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cg2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cg2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	3739	100.0	2070	US-09-821-883-7	Sequence 7, Appl1
2	3473.5	92.9	2091	US-09-821-883-9	Sequence 9, Appl1
3	2974	79.5	1665	US-09-821-883-6	Sequence 6, Appl1
4	2959.5	79.2	1692	US-09-821-883-8	Sequence 8, Appl1

5	2426	64.9	4543	9	US-09-769-508-1	Sequence 1, Appl1
6	2423	64.8	4530	9	US-09-877-177-11	Sequence 11, Appl1
7	2422	64.8	9274	10	US-09-811-123-7	Sequence 7, Appl1
8	2422	64.8	9274	10	US-09-811-115-1	Sequence 1, Appl1
9	2418	64.7	4473	9	US-09-441-411-5	Sequence 5, Appl1
10	2410.5	64.5	3768	10	US-09-811-123-8	Sequence 8, Appl1
11	2410.5	64.5	3768	10	US-09-811-115-2	Sequence 2, Appl1
12	2405.5	64.3	3768	9	US-09-854-356-9	Sequence 9, Appl1
13	2405.5	64.3	3768	9	US-09-930-125-1	Sequence 1, Appl1
14	2188.5	58.5	1437	10	US-09-821-883-10	Sequence 10, Appl1
15	1968	52.6	3935	9	US-09-854-356-10	Sequence 10, Appl1
16	1968	52.6	3935	9	US-09-870-759-117	Sequence 11, Appl1
17	1961	52.4	3771	9	US-09-854-356-11	Sequence 11, Appl1
18	1587	42.4	867	10	US-09-821-883-24	Sequence 24, Appl1
19	1199.5	32.1	1191	10	US-09-821-883-28	Sequence 28, Appl1
20	1188	31.8	1806	9	US-09-930-125-5	Sequence 5, Appl1
21	1183	31.6	1755	9	US-09-930-125-6	Sequence 6, Appl1
22	1183	31.6	1767	9	US-09-930-125-4	Sequence 4, Appl1
23	1183	31.6	1773	9	US-09-930-125-7	Sequence 7, Appl1
24	1182	31.6	651	10	US-09-821-883-26	Sequence 26, Appl1
25	1182	31.6	3537	10	US-09-821-883-30	Sequence 30, Appl1
26	803	21.5	1115	9	US-10-102-806-165	Sequence 165, Appl1
27	724.5	19.4	2643	9	US-10-172-620-15	Sequence 15, Appl1
28	724.5	19.4	3633	10	US-09-725-433-1	Sequence 1, Appl1
29	724.5	19.4	10058	9	US-09-974-298-98	Sequence 98, Appl1
30	721	19.3	5484	10	US-09-940-101-1	Sequence 1, Appl1
31	720.5	19.3	5264	10	US-09-920-300A-1731	Sequence 1731, Appl1
32	720.5	19.3	5264	12	US-10-033-528-1731	Sequence 1731, Appl1
33	714.5	19.1	4879	9	US-10-172-620-13	Sequence 13, Appl1
34	714.5	19.1	4975	10	US-09-880-107-2342	Sequence 2342, Appl1
35	713.5	19.1	2601	10	US-09-940-101-3	Sequence 3, Appl1
36	710	19.0	1958	10	US-09-867-521-1	Sequence 1, Appl1
37	684	18.3	1833	10	US-09-783-708-2	Sequence 2, Appl1
38	675	18.1	1011	12	US-10-044-090-509	Sequence 509, Appl1
39	675	18.1	1318	9	US-10-228-811-3	Sequence 3, Appl1
40	674	18.0	435	9	US-09-826-025-8	Sequence 8, Appl1
41	674	18.0	435	9	US-10-083-590-14	Sequence 14, Appl1
42	669	17.9	452	9	US-09-918-995-12765	Sequence 32765, Appl1
43	668	17.9	381	10	US-09-821-883-19	Sequence 19, Appl1
44	655.5	17.5	1046	10	US-09-925-301-390	Sequence 390, Appl1
45	610	16.3	731	12	US-10-044-090-826	Sequence 826, Appl1

## ALIGNMENTS

RESULT 1  
US-09-821-883-7  
: Sequence 7, Application US/09821883  
: Patent No. US20020061310A1  
GENERAL INFORMATION:  
APPLICANT: Laus, Rehner  
APPLICANT: Vidovic, Damir  
APPLICANT: Graddis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022-30  
CURRENT APPLICATION NUMBER: US/09/821, 883  
CURRENT FILING DATE: 2001-03-30  
PRIORITY APPLICATION NUMBER: US 60/193, 504  
PRIORITY FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7  
LENGTH: 2070  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500-HGM-CSF construct  
US-09-821-883-7  
Alignment Scores: 3.36e-315 Length: 2070  
Pred. No.: 3739.00 Matches: 690  
Score:

Percent Similarity: 100.008  
 Best Local Similarity: 100.008  
 Query Match: 100.008  
 DB: 10  
 Caps: 0

Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Caps: 0

US-09-821-883-2 (1-690) x US-09-821-883-7 (1-2070)

QY 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20  
 DB 1 ATGAGACCTGACCCCTCTCTGCGCAGGACCAAGCCTTACCTGGCTTCTTGT 60  
 QY 21 LeuLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 40  
 DB 61 CTGCTTTTCTGCTAGACCGAAGTGTACTACCAAGAGGAGTGGCGCGCGCGCG 120  
 QY 41 SerThrGlnValCysThrGlyThrAspMetLeuLeuArgLeuProAlaSerProGluThr 60  
 DB 121 TCGACCCAGAGTGCACCGGACAGACATGAGTGGCGCTCCCTGCCAGTCCGAGACC 180  
 QY 61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80  
 DB 181 CACCTGGACATGCTCCCGCCTCTTACAGGCTGCGCAGGCTGCGAAGACCTGGAA 240  
 QY 81 LeuThrTyLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100  
 DB 241 CTCACCTACCTGCCACCAATGCCAGCTGTCTCTCTGCGAGATATCCAGAGGTGCG 300  
 QY 101 GlnTyValLeuLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120  
 DB 301 GGTACAGCTGCTCAGCTGCTCAACCAACCAAGTGAAGGCTCCACAGTCCAGAGGCTCGGAT 360  
 QY 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyAlaLeuAlaValLeuAspAsnGlyAsp 140  
 DB 361 GTGGAGGACCCAGCTCTTGTGAGACACTATGCTGCGCTGCTGAGACAAATGGAGAC 420  
 QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlyLeuGln 160  
 DB 421 CCGCTGAAACAATACCACTCCCTGTACAGGGGCTCCCGAGAGGCTGCGGAGCTGCGAG 480  
 QY 161 LeuArgSerLeuThrGlnLeuLeuGlyGlyValLeuLeuGlnArgAsnProGlnLeu 180  
 DB 481 CTTCGAGAGCTCAACAGATCTTGAAGAGGGCTTGAATCCAGCGAAGCCCGACAGCTC 540  
 QY 181 CysThrLeuAsnThrThrIleLeuThrIleLeuThrIlePheHisIleAsnAsnGlnLeuAlaLeu 200  
 DB 541 TGTACACAGAGACAGATTTGTGAGAGACATCTTCCACAAAGAACCAACTGGCTCTC 600  
 QY 201 ThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLeuGly 220  
 DB 601 ACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTCTTCTCCGATGTGTAAGGC 660  
 QY 221 SerArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240  
 DB 661 TCCCGCTGCTGGGAGAGATTCTGAGAGATTGTCAAGAGCTGAGCGCACAGTGTGCGC 720  
 QY 241 GlnGlyCysAlaArgCysGlyGlyProLeuProThrAspCysHisGlnGlnCysAla 260  
 DB 721 GGTGGCTGTCCCGCTCAAGGGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 QY 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280  
 DB 781 GCCGGCTGACGGGCGCCCAAGCACTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 QY 281 GlnIleCysGlnLeuLeuHisCysProAlaLeuValThrTyAsnThrAspThrPheGlnSer 300  
 DB 841 GGCATCTGTAGCTGACCTGCCAGCCCTGTGCTACCTACCAACAGACAGACAGTGTGATCC 900  
 QY 301 MetProAsnProGluGlyArgTyThrPheGlyAlaSerCysValThrAlaCysProTy 320  
 DB 901 ATCCCAATCCAGAGGCGCGTATACATGCGCGCCAGCTGTGACTGCTGCTGCTGCTGCT 960  
 QY 321 AsnTyLeuSerThrAspValGlySerGlyAlaGlyIleMetValHisIleAsnArg 340

DB 961 AACTACCTTTCTACGACGCTGGGATCGGGCGCTGGGGCATGTCCACACAGGACCCG 1020  
 QY 341 SerSerSerThrArgSerGlyGlyIleAspLeuThrLeuGlyLeuGluProSerGluGlu 360  
 DB 1021 AGCTCATCTACAGAGATGGGGGTGGGAGCTGACACTAGGGCTGTGAAGAC 1080  
 QY 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380  
 DB 1081 GAGGCCCCAGGCTCCACATGGCAGCCCTCCCAAGGGGCTGCTCGATTTGATGAGT 1140  
 QY 381 AspLeuGlyMetGlyAlaAlaValGlyLeuGlnSerLeuProThrHisAspProSerPro 400  
 DB 1141 GACCTGGAGATGGGGGAGCCCAAGGGGCTGCAAGGCTCCACACATGAGCCAGGCT 1200  
 QY 401 LeuGlnArgTySerGlnAspProThrValProLeuProSerGlnThrAspGlyTyVal 420  
 DB 1201 CTACAGGGGTACAGTGAAGACCCACAGTACCTGCTGCTGTGAGACTGATGAGCTAGCT 1260  
 QY 421 AlaProLeuThrCysSerProGlnProGluTyValAsnGlnProAspValArgProGln 440  
 DB 1261 GCCCCCTGACCTGACCCCGCCACCTGAAATGTGAACCCAGAGATGTTGCGGCCCG 1320  
 QY 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460  
 DB 1321 CCCCCTTGGCCCGGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 QY 461 ArgAlaIleThrLeuSerProGlyIleAsnGlyValValIleAspValPheAlaPheGly 480  
 DB 1381 AGGCGCAAGACTCTCTCCCGAGAGAAATGGGCTGCAAGAGCTTTTGTGCTTGGG 1440  
 QY 481 GlnValValGlnAsnProGluTyThrLeuThrProGlnGlyAlaAlaProGlnProHis 500  
 DB 1441 GGTCCGCTGGAGAACCCCGAATCTTGAACCCAGGAGAGAGTGGCCCTCACCCAC 1500  
 QY 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyTyTrpAspGlnAspProPro 520  
 DB 1501 CTTCTCTGCTGCTCAGCCCGACCTTGGACACTCTATTTCTGGAGACAGACCCACA 1560  
 QY 521 GlnArgGlyAlaProProSerThrPheGlyGlyThrProThrAlaGlnAsnProGluTy 540  
 DB 1561 GAGCGGGGGGCTCCACAGACCTTCAAGGGGACCTTCAAGGAGAGAGTGGCCCTCAC 1620  
 QY 541 LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSerThrGlnPro 560  
 DB 1621 CTGGGCTGGAGCTGCGAGGCGCCGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
 QY 561 TrpIleHisValAsnAlaIleGlnGlnAlaArgArgLeuLeuAsnLeuSerArgAspThr 580  
 DB 1681 TGGAGCATGTGATGCTCATCCAGAGGCGCGGCTCTCTCGAATCTGATAGAGACACT 1740  
 QY 581 AlaAlaGlnLeuValGlnThrValGlnValIleSerGlnMetPheAspLeuGlnGluPro 600  
 DB 1741 GCTCTGAGATGATGAAGACAGTAGAAGTATCTCAAAATGTTTGACTTCCAGACCCG 1800  
 QY 601 ThrCysLeuGlnThrArgLeuGlnLeuTyThrIleGlnIleAsnGlySerLeuThrIle 620  
 DB 1801 ACCGCTTACAGACCCCGCTGAGCTTACAAAGAGGCTGCGGGCAGCTTACCAAG 1860  
 QY 621 LeuArgGlyProLeuThrMetMetAlaSerHisIleTyGlnHisCysProProThrPro 640  
 DB 1861 CTCAAGGCGCCCTTGAACATGATGGCCAGCCACTACAAACAGCATCTCCCTCCACCCCG 1920  
 QY 641 GluThrSerCysAlaThrGlnIleIleThrPheGlnSerPheTyGlnAsnLeuLysAsp 660  
 DB 1921 GAAACTTCCTGTGAACCCAGATATATACCTTTGAAAGTTTCAAGAAAGAACTGAAAGAC 1980  
 QY 661 PheLeuLeuValIleProPheAspCysTrpGluProValGlnGlnGlyAlaProProPro 680  
 DB 1981 TTTCTGCTTCAATCCCTTGTGACTGCTGGAGAGCACTGAGAGGCGCGGCCACCCCG 2040  
 QY 681 ProAlaAlaAlaHisHisHisHisHis 690  
 DB 2041 CCGGCGCGCGACATCACATCACATCAC 2070



Db 1621 ACACCTACGGCAGAGAACCCAGACTGCTGGACCTGCCAGCGGCCGCCACC 1680  
Qy ArgSerProSerProSerThrGlnProThrGlnHisValAsnAlaIleGlnGluAlaArg 571  
Db 1681 CGCTACACCCAAACCTGTCACCGGCGCTGAGACATGTATGCCATCAAGAAAGCTCTG 1740  
Qy 572 ArgLeuAsnLeuSerThrArgAspThrAlaIleGluMetAsnGluThrValGluValIle 591  
Db 1741 AGCCTCTTAATGACATGCTGCTGTGAGACGAAAAGAACAGACATGACATCATC 1800  
Qy 592 SerGluMetPheAspLeuGlnGluProThrCysLeuGlnThrArgLeuGluThrVal 611  
Db 1801 TCTAATGAGTTCCTCCATCCAGAGCCGACATGTGTGACAGCCGCTGAAGCATATCAAG 1860  
Qy 612 GlnGlyLeuArgGlySerLeuThrLysLeuLysGlyProLeuThrMetAlaSerHis 631  
Db 1861 CAGGCTCTACGGGCAACCTCAACCAATGATGCGCCCTTGAACCATGATAGCAGCAC 1920  
Qy 632 TyrLysGlnHisCysProProThrProGluThrSerCysAlaThrGlnIleThrPhe 651  
Db 1921 TACACAGACGAACTGCTCCACACCCGAAACGACATGTAATGAAATGACACCTTT 1980  
Qy 652 GluSerPheLysGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysThrGlu 671  
Db 1981 GAGGATTTCTAATAGAACCTTAAAGCTTCTGTTGATATCCCTTTTGACTGTGAGAG 2040  
Qy 672 ProValGlnGluAlaProProProProAlaAlaHisHisHisHisHis 690  
Db 2041 CCGGTCCAGAAAGGCGGCGCCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2091

RESULT 3  
US-09-821-883-6  
Sequence 6, Application US/09821883  
Patent No. US20020061310A1  
GENERAL INFORMATION:  
APPLICANT: Laus, Retner  
APPLICANT: Vidovic, Damir  
APPLICANT: Graddis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821, 883  
PRIORITY FILING DATE: 2001-03-30  
PRIORITY FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6

LENGTH: 1665  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500 construct  
US-09-821-883-6

Alignment Scores:  
Pred. No.: 5,94e-249 Length: 1665  
Score: 2974.00 Matches: 554  
Percent Similarity: 99.46% Conservative: 0  
Best Local Similarity: 99.46% Mismatches: 1  
Query Match: 79.54% Indels: 3  
Db: 10 Gaps: 1

US-09-821-883-2 (1-690) x US-09-821-883-6 (1-1665)

Qy 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20  
Db 1 ATGAGAGCTGACCCCTCTCTGCGCAGGCGCAAGCCTTGGCTTCTTGT 60  
Qy 21 LeuLeuPhePheThrLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40  
Db 61 CTGCTTTTTCGTGCTAGACCAAGTGTACTAGCCAGAGAGTTGGCGCGCGCGCGCG 120

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60  
Db 121 TCGACCCAAATGTGTACACCGGACAGACATGAACTGGCGCTCCCTGCCACTCCGAGACC 180  
Qy 61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnGlnValAsnLeuGlu 80  
Db 181 CACCTGGACATGCTCCGCCACCTCTTACAGAGGCTGGCAGGTGGCGGAGAAACCTGGAA 240  
Qy 81 LeuThrThrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100  
Db 241 CTCACCTACCTGGCCACCAATGCGACAGCTGTCTCTCTGAGATATCCAGAGAGTGCAG 300  
Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120  
Db 301 GCGTACCTGCTCATCGCTCAACCAAGTAGAGAGGTCCCATGTGACAGAGCTGCGGATTT 360  
Qy 121 ValArgGlyThrGlnLeuPheGluAspAspTyrAlaLeuAlaValLeuAspAsnGlyAsp 140  
Db 361 GTGGAGGACCCAGCTCTTGAAGACACTATGCTGGCGCTGTGATGACAAATGAGAGAC 420  
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeuGln 160  
Db 421 CCGGTGAACAATACACCCCTGTACAGAGGCGCTCCCGAGAGGCGCGGAGCTGCAG 480  
Qy 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180  
Db 481 CTTCGAAGCCTCACAGAGATCTTGAAAGAGGGGTCTGTGATCCAGCGGAACCCCGACCTC 540  
Qy 181 CysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200  
Db 541 TGCTACACGACACGATTTGTGTGAGAGACATCTTCACAAAGAACACAGCTGCTCTC 600  
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220  
Db 601 ACACTGATACACCAACCAACCGCTCTGCGGCGCTGCCACCCGTTCTCGAGATGTAAAGGC 660  
Qy 221 SerArgCysThrPclLysLeuSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240  
Db 661 TCCCGGCTGCGGGAGAGAGATTTGTAGAGATGTGACAGCCTGACGCGCCTGCTGCGC 720  
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 260  
Db 721 GGTGGCTGTCCCGCTGCAAGGGGCCACTGCCACTGACTGCTGCTGAGACAGTGGCT 780  
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280  
Db 781 GCGGGGCGCAGGGGCCCAAGCACTGACTGCTGCGCTGCCCTCCACTCAACACACAGT 840  
Qy 281 GlyTyrLeuCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300  
Db 841 GGCATCTGTGAGCTGCACTGCCAGCCCTGTGACCTACACACAGACAGCTTTGAGTCC 900  
Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProThr 320  
Db 901 ATGCCCAATCCGAGGCGCGGTATACATTGGCGCCAGCTGTGTACTGCTGTCCCTAC 960  
Qy 321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyLysMetValHisHisArgHisArg 340  
Db 961 AACTACTTCTTCAAGGACGTGGGATCGGGCGGTGGGGGATGATGCCACACAGCAGCAGCCG 1020  
Qy 341 SerSerSerThrArgSerGlyGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360  
Db 1021 AGCTCATCTACACAGAGGTGGGTGGGACCTGACACTAGAGGTGGGCGCTGTGAAGAG 1080  
Qy 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380  
Db 1081 GAGGCCCCCAGGTCTCTCACTGCGACCTCCGAAGGGGCTGGCTCCGATGTATTGATGCT 1140  
Qy 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400  
Db 1141 GACCTGGGAATGGGGGAGCAGCAAGGGGCTCAAGAGCTCCCAACATGACCCAGCCCT 1200  
Qy 401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420

Db 1201 CTACAGCGGTACAGTACAGACCCACACTACCCCTGAGACGATGAGCTACGTT 1260  
Qy 421 AAlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440  
Db 1261 GCGCCCTGACCTGAGCCCGCCAGCTGATATGTGAACCAAGCCGATGTTGGGCCAG 1320  
Qy 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460  
Db 1321 CCCCCCTGGCCCCGAGAGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
Qy 461 ArgAlaIleuThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480  
Db 1381 AGGCGCAAGACTCTCTCCCGAGGAGAAATGGGGCTGCAAAAGACTTTTGGCTTGGG 1440  
Qy 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHis 500  
Db 1441 GGTGCGGTGAGAACCCCGAGACTTGTACACCCCGAGGAGAGCTGCCCCCTACGCCAC 1500  
Qy 501 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro 520  
Db 1501 CCTCTCTGCTGCTTACGCGCCAGCCCTTGACAACTCTATTACTGGAGACAGACCCACA 1560  
Qy 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 540  
Db 1561 GAGCGGGGGCTCCACCCAGCAGACCTTCAAGGAGACACCTACGAGGAGAACCCAGAGTAC 1620  
Qy 541 LeuGlyLeuAspValProAlaAlaAlaArgSerProSerProSer 557  
Db 1621 CTGGGTCTGAGCTGACGAGCGGCC-----GCACA-TCACCATTCACCATCA 1664

RESULT 4  
US-09-821-883-8  
Sequence 8, Application US/09821883  
Patent No. US20020061310A1  
GENERAL INFORMATION:  
APPLICANT: Laus, Reiner  
APPLICANT: Vidovic, Damir  
APPLICANT: Gradalis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821,883  
CURRENT FILING DATE: 2001-03-30  
PRIORITY APPLICATION NUMBER: US 60/193,504  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 1692  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500\* construct  
US-09-821-883-8

Alignment Scores:  
Pred. No.: 1,1e-247 Length: 1692  
Score: 2959.50 Matches: 554  
Percent Similarity: 97.88% Conserved: 0  
Best Local Similarity: 97.88% Mismatches: 1  
Query Match: 79.15% Indels: 12  
DB: 10 Gaps: 2

US-09-821-883-2 (1-690) x US-09-821-883-8 (1-1692)

Qy 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20  
Db 1 ATGAGAGCTGACCCCTCTCTGCGCAGGCGAGCAAGCCTTAGCCTTGGCTTCTTT 60  
Qy 21 LeuLeuPhePheThrLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40  
Db 61 CTGCTTTTCTGCTGAGACCGAAGTGTACTAGCCAAAGAGATTGGCGCGGCCCGCG 120

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuAlaArgLeuProAlaSerProGluThr 60  
Db 121 TCGACCCAAAGTGTGACCGGCGACACACATGAAGCTGGGGCTCCCTCCAGTCCGAGACC 180  
Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnLysLeuGlu 80  
Db 181 CACCTGAGACATGCTCCGCGACCTTACACAGGGCTCCAGAGTGTGCAGAGAAACCTGGAA 240  
Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIleValGln 100  
Db 241 CTCACCTACCTGCGCCCAAGCCAGCTGTCTTCGACAGATATCCAGAGAGGTGCAG 300  
Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120  
Db 301 GGCTACGTGCTCATGTGCTCACACCAAGTAGAGGCTCCACTGACAGAGTGGGAGTT 360  
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140  
Db 361 GTGCGAGGACCCACACTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAAATGAGAC 420  
Qy 141 ProLeuAspAsnThrThrProValThrGlyAlaSerProGlyLysLeuArgGluLeuGln 160  
Db 421 CCGTGAAACAATACCAACCCCTGTACAGGGGCTCCCGAGAGGCTCGGGAGCTGCAG 480  
Qy 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180  
Db 481 CTTCGAAAGCTTACAGAGATCTTGAAGAGAGGGGCTTGTATCCAGGAAACCCAGCTGC 540  
Qy 181 CysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200  
Db 541 TGCTACCAAGACACGATTTTGTGAAGACATCTTCCACAGAACACACAGCTGCTCTTC 600  
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220  
Db 601 ACAGTATAGACACCAACCGCTCTGCGCTGCGACCCCTGTCTCCGATGTGTAGAGGC 660  
Qy 221 SerArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValLysAla 240  
Db 661 TCCGCTGCTGGGAGAGAGTCTGAGAGTGTCTGAGAGCTGAGCGGACCTGCTGTCG 720  
Qy 241 GlyLysCysAlaArgCysLysGlyProLeuProThrAspCysGlnHisGluLysAla 260  
Db 721 GGTGCTGTGGCCCTGCAAGGGGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Qy 261 AlaGlyCysThrGlyProLysHisSerAspLysAlaCysLeuHisPheAsnHisSer 280  
Db 781 GCGGCTGACGCGGCCCAAGACATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValIleThrAsnThrAspThrPheGluSer 300  
Db 841 GGCATCTGTGAGCTGACAGTCCGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValIleThrLysAspProTyr 320  
Db 901 ATGCCCAATCCCGAGGCGCGGTATGATTCGCGCGAGCTGTGACTGCTGCTGCTGCTG 960  
Qy 321 AsnTyrLeuSerThrAspValGlySer-----GlyAla 331  
Db 961 AACTACCTTCTTACGAGCTGGGATCCGCTACCATCATTAATTGGAGAAAGTTGGCGCT 1020  
Qy 332 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyLysLeuLeu 351  
Db 1021 GGGGCGATGCTGCACACAGGAGCGGACGCTATCTACAGAGAGTGGGCGGAGCTG 1080  
Qy 352 ThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGlu 371  
Db 1081 ACACTAGGGCTGAGCCCTCTGTAAGAGAGGCGCCAGGTCTCCACTGCGACCTCCGAA 1140  
Qy 372 GlyAlaGlySerAspValPheAspGlyAspLeuGluMetGlyAlaAlaLysGlyLeuGln 391  
Db 1141 GGGGCTGCTCCGATGTATTGTATGTGACCTGGGAATGGGGGACCAAGGGGCTGCA 1200

```

OY 392 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 411
Db 1201 AGCCCTCCACACGATGACCCGCTCTACAGGGTACGTAGAGACCCACACTACC 1260
OY 412 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 431
Db 1261 CTGGCTCTGAGACTGATGGCTACCTTCCCTGACCTGACGCCCCACAGCTGAAATAT 1320
OY 432 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 451
Db 1321 GTGAACCGACGAGATGTTCCGCGCCAGCCCTTCGCCCCAGAGAGGGCTCTGCTGCT 1380
OY 452 AlaArgProAlaGlyAlaThrLeuGlnArgAlaGlyThrLeuSerProGlyLysAsnGly 471
Db 1381 GCCGACCTGCTGTGGCTGACCTGAAAGGGCCAAAGACTCTCTCCCGGAGAAAGATGG 1440
OY 472 ValValLysAspValPheAlaPheGlyGlyValAlaGluAsnProGluTyrLeuThrPro 491
Db 1441 GTCTCAAAAGACGTTTGGCTTTGGGGGTGCTGCGTGGAGAACCCGAGTACTTGACACCC 1500
OY 492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511
Db 1501 CAGGAGAGAGCTGCCCTCACCCCTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
OY 512 LeuTyrTyrTyrPaspGlnAspProProGluArgGlyAlaProProSerThrPheLysGly 531
Db 1561 CTCTATTACTGTGGACCGAGACCCACAGAGGGGGGCTCCACCCAGCACCTTCAAAAGG 1620
OY 532 ThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaProAla 551
Db 1621 ACACCTACGGAGAGAACCCAGACTACTGCTGGTCTGGACGTGCCAGCGGCC-----GCA 1674
OY 552 ArgSerProSerProSer 557
Db 1675 CA-TCACCATCACCATCA 1691

RESULT 5
US-09-769-508-1
: Sequence 1, Application US/09769508
: Patent No. US2002015527A1
: GENERAL INFORMATION:
: APPLICANT: STUART, SUSAN G.
: APPLICANT: MONAHAN, JOHN J.
: APPLICANT: LANGTON, BEATRICE CLAUDIA
: APPLICANT: HANCOCK, MIRIAM E.C.
: APPLICANT: BLUFORD, PETER
: TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
: FILE REFERENCE: BBIO-111-C1
: CURRENT APPLICATION NUMBER: US/09/769,508
: CURRENT FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4543
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (150)..(3914)
US-09-769-508-1

Alignment Scores:
Pred. No.: 7,1e-201 Length: 4543
Score: 2426.00 Matches: 525
Percent Similarity: 40.66% Conservative: 2
Best Local Similarity: 40.51% Mismatches: 23
Query Match: 64.88% Indels: 746
DB: 9 Gaps: 4

US-09-821-883-2 (1-690) x US-09-769-508-1 (1-4543)
OY 2 ArgAlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPhe----- 18

```

```

Db 63 CGGCCCCACCCCTTCGACGACACCCCGCGCCCTCCAGCCGGGTCCAGCCGAGC 122
OY 19 -----LeuPheLeuLeuPheThrLeuAsp 27
Db 123 CATGGGGCGGAGCCGACGTAGACACATGGAGCTGGGGGCTTGTGCCCTGG----- 176
OY 28 ArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThrGlnValCysThrGly 47
Db 177 ---GGGCTCTCTCCGCTCTTGGCCCTTGGCCCGGAGCGCGACACCCAGATGCTCGCGGC 233
OY 48 ThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeuAspMetLeuArgHis 67
Db 234 ACAGACATGAAGCTCGGCTGCTCCAGTCCCGAGCCACCTGGACATGCTCGCGCAC 293
OY 68 LeuTyrGlnGlyCysGlnValValGlnLysAsnLeuGluLeuThrTyrLeuProThrAsn 87
Db 294 CTCTACCGAGGCTGCGAGGTGGTGCAGGAAACCTGGAACTCCTACCTGCCACCAAT 353
OY 88 AlaSerLeuSerPheLeuGlnAspIleGlnLysValGlnGlyTyrValLeuIleAlaHis 107
Db 354 GCCAGCTGTCTCTTCCTGACAGATATCCAGAGGTGACAGGGCTACGTCTCATGCTCAC 413
OY 108 AsnGlnValArgGlnValProLeuGlnArgLeuArgIleValArgGlyThrGlnLeuPhe 127
Db 414 AACCAAGTGAAGGACAGTCCACCTGACAGAGGTGGGATTTGGAGGACCCAGCTCTT 473
OY 128 GluAspAsnTyrAlaLeuAlaValLeuAspAsnGlnAspProLeuAspAsnThrThrPro 147
Db 474 GAGGACAACTATGCCCTGGCCGCTGCTGACAAATGAGACCCGCTGACAAATACCACTT 533
OY 148 ValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIle 167
Db 534 GTCAAGAGGGCTCCCGCCAGAGGCTCGGGAGGTGAGCTTCCAAAGCTCACAAGAGATC 593
OY 168 LeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeu 187
Db 594 TTGAAGAGGGGTCTTGATCCAGGGGAACCCAGCTCTCTACCAAGACAGCATTTTG 653
OY 188 TrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArg 207
Db 654 TGAAGAGCATCTTCCACAAAGAACACAGCTGGCTGCACATGATGACACCAACCGC 713
OY 208 SerArgLacysHisProCysSerProMetCysLysGlySerArgCysTyrGlyLysIle 227
Db 714 TCTGGGGCTGCCACCCCTGTTCTCCGATGTGAAGGGCTCCCGCTGCTGGGAGAGAGT 773
OY 228 SerGluAspCysGlnSerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLys 247
Db 774 TCTAGGATTTCTCAGAGCTGACGGCGACGTGTGTGCGGGTGTGCGCGCTGCAGAG 833
OY 248 GlyProLeuProThrAspCysCysHisGlnGlnCysAlaAlaGlyCysThrGlyProLys 267
Db 834 GGCGCAGTCCCACTGACTGCTGTCATGAGAGAGTGTGCTGGCTGCAGCGGCCCAAG 893
OY 268 HisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyIleCysGlnLeuHisCys 287
Db 894 CACTGTGACTGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 953
OY 288 ProAlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAsnProGluLysArg 307
Db 954 CCAGCCCTGTGCTCAACACAGACACGTTTAGTCCATGCCCCAATGCCGAGGGCGCG 1013
OY 308 TyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspVal 327
Db 1014 TATACATTTGGCGGCGACACTGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
OY 328 GlySer----- 329
Db 1074 GGATCTTGACACCTGCTGTGCCCCCTGCACAAACCAAGAGTGACAGAGAGATGAACA 1133
OY 329 ----- 329

```



Dd 1134 CAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCAGTGTCTATGTCTGGGCATGGAG 1193  
Oy 329 ----- 329  
Dd 1194 CACTTGGCAGAGGTGAGGGCAGTACAGTGCCTAATTCAGAGAGTTGTGGCTGCAAG 1253  
Oy 329 ----- 329  
Dd 1254 AAGATCTTTGGAGGCTGGCAATTTGTGCCGGAGACTTTGATGGGGACCCAGCCTCCAAAC 1313  
Oy 329 ----- 329  
Dd 1314 ACTGCCCGCTCCAGCCAGAGCAGCTCCAAAGTTTGTAGACTGTGGAAGATCACAGGT 1373  
Oy 329 ----- 329  
Dd 1374 TACCTATACATCTCAGCATGGCCGGACAGCCTGCCTGACCTCAGCGTCTTCCAGAACCTG 1433  
Oy 329 ----- 329  
Dd 1434 CAAGTAATCCGGGGAGCAATTTGTCAATATGGCGCCTACTGCTGACCCGTGCAAGGGCTG 1493  
Oy 329 ----- 329  
Dd 1494 GGCATCAGCTGGCTGGGCTGCGCTCAGTGAAGGAATGGCGAGTGAAGTGGCCCTCAATC 1553  
Oy 329 ----- 329  
Dd 1554 CACCAATACACCCACCTCTGCTTGTGTCACACAGGTGCCCTGGGACACAGCTCTTTGGGAAC 1613  
Oy 329 ----- 329  
Dd 1614 CCGCACAAGCTGTGCTCCACACTGCAACCGGCCAGAGAGAGATGTGTGGCGAGGCG 1673  
Oy 329 ----- 329  
Dd 1674 CTGGCTGCCACACAGCTGTGCGCCGCAAGGCAGCTGTGGGTCAAGGCCACCCAGTGT 1733  
Oy 329 ----- 329  
Dd 1734 GTCAACTGCAGCAGTTCCTTGGGGCCAGAGTGCCTGGAGGAATGCCAGTACTGCAG 1793  
Oy 329 ----- 329  
Dd 1794 GGGCTCCCAAGGAGATGTGAATGCAGGCACTGTTGCCGTGCACCCCTGATGTGAG 1853  
Oy 329 ----- 329  
Dd 1854 CCCCAGAAATGGCTCAGTACCTGTTTGGACCGAGGCTGACAGTGTGTGGCTGTGCG 1913  
Oy 329 ----- 329  
Dd 1914 CACTAATAGGACCTCCCTTTCGTGGCGCCGCTGCCAGCGGTGTGAACCTGACCTC 1973  
Oy 329 ----- 329  
Dd 1974 TCCTATATGCCCATCTGGAAGTTTCCAGATGAGGAGGGGCATGCGACGCTTGGCCCATC 2033  
Oy 329 ----- 329  
Dd 2034 AACTGACCACTCTGTGTGTGACCTGTGATGACAAGGGCTGCCCGCGAGAGAGAGCC 2093  
Oy 329 ----- 329  
Dd 2094 AGCCCTGTGAGCTCCATGCTCTGTCCGGGTGTGGCAATTTGCTGTGCTGTGTTGGGG 2153  
Oy 329 ----- 329  
Dd 2154 GTGGCTTTGGGATCTCATCAAGGAGCGGACGACGAAGATCCGAAGTACACGATGGCG 2213  
Oy 329 ----- 329  
Dd 2214 AGACTGTGCAAGAAACGAGCTGTGTGAGCGGCTGACACTAGCGAGCGATGCCAAC 2273

Oy 329 ----- 329  
Dd 2274 CAGCGCAGATGCGGATCTCTGAAGAGACGAGAGCTGAGGAAGGTGAAGCTTGATCT 2333  
Oy 329 ----- 329  
Dd 2334 GGGCTTTTGGCAGCTCTACAGGGCATGTGATCCCTGATGGGAGAAATGTGAATAAT 2393  
Oy 329 ----- 329  
Dd 2394 CCAGTGGCCATCAAGTGTGAGGAAAAACACATCCCAAGGCAACAAAGAAATCTTA 2453  
Oy 329 ----- 329  
Dd 2454 GACGAAGCATACGTATGCTGTGTGGGCTCCCATATGTCCTCCGCTTCTGGGCATC 2513  
Oy 329 ----- 329  
Dd 2514 TGCTGACATCCAGGTCAGCTGTGTGACAGACTTATGCCCTATAGCTGCTTTAGAC 2573  
Oy 329 ----- 329  
Dd 2574 CATGTCCGGAAAAACCGGAGCGCTGGGGCTCCAGAGCTGTGAAGTGTATGACAG 2633  
Oy 329 ----- 329  
Dd 2634 ATTGCCAAGGGATGAGCTACTGAGGAGATGTGGCGCTGTACACAGGAGACTTGGCGCT 2693  
Oy 329 ----- 329  
Dd 2694 CGAAGCTGTGTGTCAGATGCCAACCATGTCAAAATTACAGACTTGGGCTGGCTGG 2753  
Oy 329 ----- 329  
Dd 2754 CTGTGACATTTGACAGACAGAGTACATGAGTGGGGCAAGGTGCCCTCAAGTGG 2813  
Oy 329 ----- 329  
Dd 2814 ATGGGCTGTGAGTCCATTTCCGGCGGGGTTCCACCACAGAGTGTGTGAGTTAT 2873  
Oy 329 ----- 329  
Dd 2874 GGTGTACTGTGTGGAGCTGATGACTTTTGGGGCCAAACCTTACGATGATCCAGCC 2933  
Oy 329 ----- 329  
Dd 2934 CCGGAGATCCCTGACCTGTGAAAAAGGGGAGCGGCTGCCAGCCCCCATCTGTAC 2993  
Oy 329 ----- 329  
Dd 2994 ATTGATGCTACATGATCATGTGCAATGTGTGATGATGACTTGAAATGTGGGCCAAG 3053  
Oy 329 ----- 329  
Dd 3054 TTCGGGAGTTGTGTCTGAATTTCTCCGCATGCCAGGAGACCCAGCGCTTGTGTGTC 3113  
Oy 329 ----- 329  
Dd 3114 ATCCAGAAATGAGACTTGGGCCAGCCAGTCCCTTGGACAGCACTGTACCGTCACTG 3173  
Oy 329 ----- 329  
Dd 3174 CTGAGAGCATGACATGGGGGACCTGTGTGATGCTGAGAGATATCTGTATCCCAAGCAG 3233  
Oy 329 ----- 329  
Dd 3234 GGGCTCTTGTGTCCAGACCTGCCCCGGCGGTGGGGCATGTGTCCACACAGGACCGC 3293  
Oy 329 ----- 329  
Dd 3294 SerSerSerThrArgSerGlyGlyLysPleuThrLeuGlyLeuGluProSerGluLys 360  
AGCTCATCTACAGAGAGTGGGGGTGGGACCTGACACTAGGGGCTGTGAGCCCTCTGAAAGG 3353



Db 1225 GCCAATATCAGAGAGTTGCTGCTGCAAGAGATCTTTGGAGCGCTGCAATTTCTGCCG 1284  
Oy 329 ----- 329  
Db 1285 GAGAGCTTTGATGGGAGCCAGCCTCCAACTGCTCCCGCTCCAGCAGAGCACTCCAA 1344  
Oy 329 ----- 329  
Db 1345 GTTTGAGACTCTGGAGAGATCAGAGTTACTATATCATCTCAGCATGGCCGGAGAGC 1404  
Oy 329 ----- 329  
Db 1405 CTGCTGACTCAGCGCTCTCCAGAACCTGCAGATATCCGGGAGCAATTTCTGCACAA 1464  
Oy 329 ----- 329  
Db 1465 GCGGCTTACTGCTGACCCCTGCAGAGGCTGGGCATCAGCTGGCTGGGCTCGCTCAGCTG 1524  
Oy 329 ----- 329  
Db 1525 AGGGAATGGGACATGAGACTGGCCCTCATCCACATTAACACCACCTCTGCTTCTGCAC 1584  
Oy 329 ----- 329  
Db 1585 ACGTGCCCTGGGAGCAGCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAAC 1644  
Oy 329 ----- 329  
Db 1645 CGGCGAGAGAGAGTGTGTGGGCGAGGCGCTGGCCCTGCACAGCTGTGCCCGAGAG 1704  
Oy 329 ----- 329  
Db 1705 CACTGCTGGGGTCCAGGGCCCAACAGTGTGTCAACTGCAGCAGTTCTTGGGGCCAG 1764  
Oy 329 ----- 329  
Db 1765 GAGTCGTGAGAGATGCCGAGTACTGCAGGGGCTCCGAGGAGATGTGAATGCCAGG 1824  
Oy 329 ----- 329  
Db 1825 CACTGTTCGCGTGCACCTGAGTGTAGCCCCAGAAATGGCTCAGTCACTGTTTGA 1884  
Oy 329 ----- 329  
Db 1885 CCGAGGCTGACCACTGTGTGCTGTGCCCACTATTAAGAACCTTCTTGTGCGTGGCC 1944  
Oy 329 ----- 329  
Db 1945 CGCTGCCCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATGTGGAATTTCCAGAT 2004  
Oy 329 ----- 329  
Db 2005 GAGAGGGCGCATGCGACCTTGCCCATCACTGACACCACCTCTGTGTGAGACCTGGAT 2064  
Oy 329 ----- 329  
Db 2065 GACAGGGGCTGCCCGCGAGAGAGACCAACCTCTGACCTCATGCTCTCTGCGGTG 2124  
Oy 329 ----- 329  
Db 2125 GTTGGCATTTCTGCTGTGCTGTGGGGGTGTCTTTGGGATCCTCATCAAGCAGCG 2184  
Oy 329 ----- 329  
Db 2185 CAGCAGAAAGATCCGGAAGTACAGATGCGGAGACTGTGCAGAAACGAGAGCTGTGAG 2244  
Oy 329 ----- 329  
Db 2245 CCGGTGACACCTAAGGGAGCGATGGCCAAACGAGGCGAGATCGGATCTCTGAAGAAGAG 2304  
Oy 329 ----- 329

Db 2305 GAGCTGAGAGAGTGAAGTGTGATGTGCGGCTTTTGGCACAGTCTACAGGGCATC 2364  
Oy 329 ----- 329  
Db 2365 TGGATCCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAGTGTGAGGGAAC 2424  
Oy 329 ----- 329  
Db 2425 ACATCCCCAAGCCAAAGAAATCTTAGACGAGATAGTATGGTGTGTGGC 2484  
Oy 329 ----- 329  
Db 2485 TCCCATATGTCTCCGCTTTCTGGGATCTGCTGACATCCAGCTCAGCTGTGACA 2544  
Oy 329 ----- 329  
Db 2545 CAGTTATGCCCTATGGCTGCCCTTAGAACATGTCCGGAAACCGGAGCGCTGGGC 2604  
Oy 329 ----- 329  
Db 2605 TCCAGACCTGCTGACTGTGTATGCAGATTTGCCAAGGGGATGAGCTACCTGAGAGAT 2664  
Oy 329 ----- 329  
Db 2665 GTGCGCTGTACACAGGAGCTTGGCGCTCGGAACGTGTGTCAAGATCCCAACCAT 2724  
Oy 329 ----- 329  
Db 2725 GTCAAAATTAACAGACTTGGGCTGGCTGCTGCTGTGACATTGACGAGACAGATACAT 2784  
Oy 329 ----- 329  
Db 2785 GCAGATGGGGGCAAGTGGCCATCAAGTGTGATGGGCTGTGATCCATTCTCCGCGGGG 2844  
Oy 329 ----- 329  
Db 2845 TTCACCCACAGATGATGTGTGAGTATGTGTGACTGTGTGGAGCTGATGACTTTT 2904  
Oy 329 ----- 329  
Db 2905 GGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCTGACCTCTGGAAGAGGG 2964  
Oy 329 ----- 329  
Db 2965 GAGCGGCTGCCCGACCCCACTTGCACCAATTGATGTCTACATGATCATGTCAATGT 3024  
Oy 329 ----- 329  
Db 3025 TGGATGATTACTGTGAATGTGCGCCAAAGATTCGGGATGTGTGAATTTCCCGC 3084  
Oy 329 ----- 329  
Db 3085 ATGCCAGGAGCCCCAGCGCTTGTGTATCCAGAAATGAGACTTGGGCCAGCCAGCT 3144  
Oy 329 ----- 329  
Db 3145 CCGTTGAGACGACCTTCAACCGTCACTGCTGAGAGAGATGACATGGGGAGCTGGTG 3204  
Oy 330 ----- 330  
Db 3205 GATGCTGAGAGATATCTGTGTAACCCAGCAGGCTTCTTCTGCAGACCTGCCCGGGC 3264  
Oy 331 AAGTGTGAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAG 3324  
Db 3265 GCTGGGGCATGTGTCCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3324  
Oy 331 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 370  
Db 3325 CTGACACTAGGGCTGAGCGCTCTGAAGAGAGGCCCCAGAGTCTCCACTGACACCTCC 3384  
Oy 371 GIUGIYALAGLYSERASPVALPHENASPLIYASPLEUGLYMETGLIALALALYSELYLEU 390  
Db 3385 GAAAGGGCTGGCTCCGATGTATTTGATGTGTGACCTGTGGAAATGGGGGAGCCAAAGGGCTG 3444

Qy	391	glnserleuprothrhisaspprosesproleuglnarysergluasprothrval	410
Db	3445	caaaagctcccccacacatgacccccacacctcttcaacgggtaacatgaaagacccacacata	3504
Qy	411	proleuprosesergluthraspdlglttgvallalaiproleuthrcysaserproglu	430
Db	3505	ccccctgacctgtgcagactgattggctacgtttggcccccctgacctgcagccccccagcctgaa	3564
Qy	431	tyralasnoinproasprvalatargproginproproserproatrglulglyproleupro	450
Db	3555	tatgtgaaccacagccagatgtgtcgccccaccccccttcgcccccgaaagggcccttgacct	3624
Qy	451	alaalaargproalacglualatrlhrlieuglatalatylsthrleuaserproglulysasn	470
Db	3625	gctccccgacctgctgctgctccactctgaaaaagggccaaagactctctcccccaggaagat	3684
Qy	471	glyalvaliylusasprvalrhealarphegllyglialvalgluasproglutylleuthr	490
Db	3665	gggggtcgtcaaaagcgtttttggctttggggggcgttgaaamccccgaagacttgaca	3744
Qy	491	proglnglvglyalaalarpoglnprohlsiproproglolapheserproalaphasp	510
Db	3745	ccccagagagagagctggcccttcagaccscacccgtccgttcgcttcagccacaccttgac	3804
Qy	511	asnlleuylrtytrtparsrglnasprproprogluatyglialarproproserthrhelys	530
Db	3805	aaacctgattactgaggacccagaccacacagacagacgggggggtccaccacagcactttaa	3864
Qy	531	glythrprothralsgluasnproglutyltyrleuglyleuasprvalproalalaalapro	550
Db	3865	ggagacacctacggcagagaaacccagagatcctgggtcgtgacgtgcac-----gtgtga	3918
Qy	551	alaargserproserpro	556
Db	3919	accgaagaagccaaagctccg	3936

## RESULT 7

```

Sequence 7, Application US/09811123
Patent No. US20020001587A1
GENERAL INFORMATION:
APPLICANT: Sharon Erickson
APPLICANT: Ralph Schwall
APPLICANT: Mark Slikowski
TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-erbB
TITLE OF INVENTION: ANTIBODY-MAYTANSINOID CONJUGATES
FILE REFERENCE: GENENT.073A2
CURRENT APPLICATION NUMBER: US/09/811,123
CURRENT FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/238,327
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 09/602,530
PRIOR FILING DATE: 2000-06-23
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 9274
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Vector Sequence
US-09-811-123-7

```

### Alignment Scores:

Pred. NO.:	3,81e+200	Length:	927
Score:	2422.00	Matches:	528
Percent Similarity:	41.12%	Conservative:	9
Best Local Similarity:	40.43%	Mismatches:	31
Query Match:	64.78%	Indels:	738
DB:	10	Gaps:	6

US-09-821-883-2 (1-690) x US-09-811-123-7 (1-9274)

Oy	2	ArgAlaIaIaProLeuLeuLeuAlaIaArgAlaIaAlaSerLeuSerLeuGlyPheLeuPheLeu	21
Db	1655	CGCCCCGGGGCCCTCCACGGCCGGGTCCAGCGGAGCGGAGCCATGAGCTGGC-----GCC	1745
Oy	22	LeuPhePheTrpLeuAspArgSerValLeuAlaValGlyLeuAlaArgGlyAlaAlaSer	41
Db	1746	TTTGCGCGCTGG-----GGAGCTCTCTCTCGCCCTCTTACGCCCCCGGAGCCCGAGAC	1796
Oy	42	ThrGlnValCysThrGlyTrpAspMetLeuArgLeuProAlaSerProGluTrpHis	61
Db	1797	ACCCAAAGTGTACACGGGCACAGACATGAACTCGGGCTCCGTCCAGTCCGAGACCCAC	1856
Oy	62	LeuAspMetLeuArgHisLeuTyrgInGlyCysGlnValValGlnGlyAsnLeuGluLeu	81
Db	1857	CTGGACATGCTGCCGCCACCTCTACAGAGGCTGGCAGAGTGTGTCAAGAAACCTGGAACTC	1916
Oy	82	ThrTyrlLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIleValGlnGly	101
Db	1917	ACCTACCTGGCCACACATGTCACACTGTCTCTCTCCAGAAATCCAGAGGTCCAGCGC	1976
Oy	102	TyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIleVal	121
Db	1977	TACGTGTCACTCGTCCAAACCAAGTGAAGAGGTCCCACTGCAGAGGTGGGAGTTGTG	2036
Oy	122	ArgGlyThrGlnLeuPheGlnAspAsnTyrrAlaLeuAlaValLeuAspAsnGlyAspPro	141
Db	2037	CGAGGCAACCCAGCTTTGTGAGGACAACTGACCTGGCCGTCTGACATGTGAACACCG	2096
Oy	142	LeuAsnAsnTrpTrpProValTrpGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu	161
Db	2097	CTGAACCAATACACACCTGTGCACAGGGGCTCCCAAGAGGCTCGGAGGTCCAGCTT	2156
Oy	162	ArgSerLeuThrGlnIleLeuGlyGlyValLeuIleGlnIArgAspProGlnLeuCys	181
Db	2157	CGAAGCCCTCACAGAGATCTTGAAAGGGGGGTGTGATCCAGCGAAACCCGCACTGTGC	2216
Oy	182	TyrGlnAspTrpIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr	201
Db	2217	TACCAAGCAACGATTTGTGGAAAGCACTTCCACAAACAAACCAAGCTGGCTTCACA	2276
Oy	202	LeuIleAspTrpAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer	221
Db	2277	CTGATACAGACACCAACCGCTCTCGGGGCTGCACACCTGTCTTCGCGATGTGAAGGCTTC	2336
Oy	222	ArgCysTrpGlyLeuSerSerGlnAspCysGlnSerLeuThrArgThrValCysAlaGly	241
Db	2337	CGGTCTGGGGAGAGACTTGTGAAGATTGTGCACAGGCTCAGCCGCACTGTCTGTGCCGCT	2396
Oy	242	GlyCysAlaIaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAlaIa	261
Db	2397	GGCTGTGCGGCTGCACAGGGGCCACTGTGCCACTGACTGTGCCATGAGCAGTGTCTGCC	2456
Oy	262	GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly	281
Db	2457	GGGTGACAGGGGCCCAAGCACTGTGACTGCTGGCTGGCTGCCTCAACTCAACACAGTGGC	2516
Oy	282	IleCysGlnLeuHisCysProAlaIaLeuValThrTyrrAsnTrpAspThrPheGluSerMet	301
Db	2517	ATCTGTAGCTGCATCTGCCACGCTGGGTCCACTTCAACAAACAGACAGCTTTGATGCTCATG	2576
Oy	302	ProAsnProGluGlyArgTyrrTrpPheGlyAlaSerCysValaThrAlaCysProTyrrAsn	321
Db	2577	CCCAATCCCAAGGGCCGGTATACATTCTGGGCGCAGGTGTGTACGTGCGCTGTCCCTACAAAC	2636
Oy	322	TyrLeuSerThrAspValGlySer-----	329
Db	2637	TACCTTTCTACGACGACGTGGGATCTCTGCACCTCTGTCCTGCCCCCTGCACAAACCAAGCTGC	2696
Oy	329	-----	329
Db	2697	ACAGACAGAGATGCACACAGCGGTGTGAAGATGCAGCAAGCCCTGTGGCCGAGTGTGC	2756

QY 329 ----- 329  
Db 2757 TATGGCTGGGATGAGACCTTGGAGAGGTGAGGCGATTACCAATCCATATCCAG 2816  
QY 329 ----- 329  
Db 2817 GAGTTGTGCTGCAAGAAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGCTTTGAT 2876  
QY 329 ----- 329  
Db 2877 GGGGAGCCAGCTCCAGACACTGCCCCGCTCAGCCAGACAGCTCCAGTGTGTGAGACT 2936  
QY 329 ----- 329  
Db 2937 CTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGGCCGGACAGCTGCTGACCTC 2996  
QY 329 ----- 329  
Db 2997 AGCGTCTTCAGAACCTCGAAGTAATCCGGGAGCGAATTCTGCACAAATGGCGCTACTCG 3056  
QY 329 ----- 329  
Db 3057 CTGACCTCTGCAAGGCTGGGCAATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACCTGGGC 3116  
QY 329 ----- 329  
Db 3117 AGTGGACTGGCCCTCATCCACATAAACACCACCTCTGCTTGTGTGACACAGGTGCCCTGG 3176  
QY 329 ----- 329  
Db 3177 GACCAAGCTCTTCGGAACCCGACCAAGCTCTGTCTCACACTGCCAACCAGCCAGAGAGAC 3236  
QY 329 ----- 329  
Db 3237 GAGTGTGTGGGGAGGGGCTGGCCTGCCACCAAGCTGTGCGCCGAGGGGCACTGTGGGT 3296  
QY 329 ----- 329  
Db 3297 CCAGGGCCCAACCCAGTGTCTCAACTGCAAGCCAGATTCTCTCGGGGCGAAGAGTGGCTGAG 3356  
QY 329 ----- 329  
Db 3357 GAATGCCAGTACTGCAAGGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGTTGGCG 3416  
QY 329 ----- 329  
Db 3417 TGCACACCTGAGTGTCAAGCCCCAGAAATGGCTAGTACCTGTTTGGACCGAGGCTGAC 3476  
QY 329 ----- 329  
Db 3477 CAGTGTGTGGCTGTGCCACTATAAGACCTCCCTTGTGGGTGGCCGCTGCCACAG 3536  
QY 329 ----- 329  
Db 3537 GGTGTGAACCTGACCTCTCTACATGCCCAATGTGAAATTCAGAGATGAGAGGGCGCA 3596  
QY 329 ----- 329  
Db 3597 TGCACACCTTGGCCCACTCAACTGCAACCCACTCTGTGTGGACCTGAGTACAAAGGGCTGC 3656  
QY 329 ----- 329  
Db 3657 CCGGCGAGACAGAGAGCCAGCCCTGTGAGTCCATGTGTCTGGGGTGTGGCAATTG 3716  
QY 329 ----- 329  
Db 3717 CTGTGTGTGTCTTGGGGGTGTCTTTGGATCCTCATCAAGCAGAGGAGAGAGATC 3776  
QY 329 ----- 329  
Db 3777 CGAAGTACAGATGCGGAGACTGTCTCAGAAACGGAAGCTGGTGGAGCCGCTGACACCT 3836  
QY 329 ----- 329

Db 3837 AGCGAGGATGCCCAACAGGCGCAGATGGGATCTGAAAGACGAGAGCTGAGAG 3896  
QY 329 ----- 329  
Db 3897 GTGAAGTGTCTTGATCTGGCGCTTTGGACAGTCTACAAAGGCACTGTGATCCCTGAT 3956  
QY 329 ----- 329  
Db 3957 GGGGAAATGTGAAATTCAGTGCCATCAAAAGTGTGAGGAAACACATCCCCAAA 4016  
QY 329 ----- 329  
Db 4017 GCCAACAAGAAATCTTAGACAGCAATACGATGGCTGTGTGGCTCCCATATGTGC 4076  
QY 329 ----- 329  
Db 4077 TCCGCTCTTGGGATCTGCTGACATCCAGGTGCAGCTGTGTGACACAGCTTATGCC 4136  
QY 329 ----- 329  
Db 4137 TATGGCTGCTCTTAGACATGTCCGGGAAACCCGAGCGCTGGCTCCAGAGACTCG 4196  
QY 329 ----- 329  
Db 4197 CTGAAGTGTATGCAATGTCGAAGGGATGAGCTACCTGAGATGTGGCGCTGTA 4256  
QY 329 ----- 329  
Db 4257 CACAGGACTTGGCGGCTCGAAGCTGTGTCAAGAGTCCCAACATGTCAAAATTACA 4316  
QY 329 ----- 329  
Db 4317 GACTTCGGGCTGGCTGCGCTGTCGACATTCAGAGACAGAGTACCATGCAATGGGGC 4376  
QY 329 ----- 329  
Db 4377 AAGTGCCCATCAAGTGAATGGCGCTGAGTCCATTCTCCGGCGGTTCACCCACAG 4436  
QY 329 ----- 329  
Db 4437 AGTATGTGTGAGATTAGTGTGTGACTGTGTGGAGCTGANTGACTTTGGGCCAAACT 4496  
QY 329 ----- 329  
Db 4497 TACGATGGATCCAGCCCGGAGATCCCTGACTGTCTGGAAGGAGGAGCGGCTGCC 4556  
QY 329 ----- 329  
Db 4557 CAGCCCCCATCTGCACCATGTATGTCTACATGATCATGTCAAAATGTGATATTGAC 4616  
QY 329 ----- 329  
Db 4617 TCTGAATGTGGCCAAAGATTCCGGAGTGTGTGTAATTCGCCGATGGCCAGGAG 4676  
QY 329 ----- 329  
Db 4677 CCCCAGCGTTTGTGTATCCAGAATGAGACTTGGGCCAGCAGTCCCTTGGACAGC 4736  
QY 329 ----- 329  
Db 4737 ACCTTCTACGCTCACTGCTGAGAGCATGACATGGGGAGCTGTGATGCTGAGAG 4796  
QY 330 ----- 334  
Db 4797 TATCTGTACCCACAGGCGCTTCTGTCTCAGACCTCGCCCGGCGCTGGGGCCAG 4856  
QY 335 ValHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThrLeuGly 354  
Db 4857 GTCCACACAGGACCGCAAGCTCATCTACAGAGATGGGGGTGGGACTGACACTAGG 4916  
QY 355 LeuGluProSerGluGluIuLaProArgSerProLeuAlaProSerGluGlyAlaGly 374  
|||||



Db 2637 TACCTTTACGAGCTGGGATCTGCACCTCGTCTGCCCTGCACAAACAGAGTG 2696  
QY 329 ----- 329  
Db 2697 ACAGCAGAGATGAAACACAGCGGTGTGAGAAATGCAGCAAGCCCTGTGCCGAGTGTGC 2756  
QY 329 ----- 329  
Db 2757 TATGCTCTGGGATGAGACCTTGGAGAGGTGAGGGCAGTTACCAATATATCCAG 2816  
QY 329 ----- 329  
Db 2817 GAGTTGTGCTGCAGAAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGCTTTGAT 2876  
QY 329 ----- 329  
Db 2877 GGGGACCAAGCCTCCACACTGCCCCGCTCCAGCCAGCAGCAGCTCCAAAGTTTGAAGCT 2936  
QY 329 ----- 329  
Db 2937 CTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGGCCGGACAGCCTGGCTGACCTC 2996  
QY 329 ----- 329  
Db 2997 AGCGTCTTCAGAACTGCAGTAATCCGGGAGAGAAATCTGCACAAATGGCGCTACTCG 3056  
QY 329 ----- 329  
Db 3057 CTGACCTCTGCAAGGGCTGGGATCAGCTGGCTGGGGCTGGCTCAGTGAAGGAAGTGGGC 3116  
QY 329 ----- 329  
Db 3117 AGTGAAGCTGGCCCTCATCCACCATAAACCAACCTCTGCTTGGTGACACGGTGCCCTGC 3176  
QY 329 ----- 329  
Db 3177 GACCACTCTTTGGAAGCCGACCAAGCTCTGCTCCACACTGCCAAACCGCCAGAGAGAC 3236  
QY 329 ----- 329  
Db 3237 GAGTGTGGGGGAGGGCCTGGCCTGCCACAGCTGTGGCCCGAGAGGACATGCTGGGT 3296  
QY 329 ----- 329  
Db 3297 CCAGGGCCCAACCAAGTGTGCAACTGCAGCAGTTCTTCGGGGCCAGAGTCCGTGGAG 3356  
QY 329 ----- 329  
Db 3357 GAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGAAATGCCAGGACATGTTTGCCG 3416  
QY 329 ----- 329  
Db 3417 TGCCACCTGAGTGTGACGCCCAAGATGGCTCAGTACCTGTTTGGACCGAGGCTGAC 3476  
QY 329 ----- 329  
Db 3477 CAGTGTGGCCCTGTGCCACTATAGAACCTCCCTCTGCGGTGCCGCTGCCAGC 3536  
QY 329 ----- 329  
Db 3537 GGTGTGAAACCTGACCTCTCTACATGCCATCTGGAAGTTCCAGATGAGAGGAGCGCA 3596  
QY 329 ----- 329  
Db 3597 TGCCAGCTTTGCCCATCAACTGCACCCACTCTGTGTGAGCCTGATGACAAAGGCTGC 3656  
QY 329 ----- 329  
Db 3657 CCGCGCAGACAGAGCAGACCCCTCTGAGCTCATGCTCTGCGGTGGATTGCAATTCTG 3716  
QY 329 ----- 329  
Db 3717 CTGCTGTGTGCTTTGGGGGTGTCTTTGGGATCCTCATCAAGCGAGCGACAGAAAGATC 3776

QY 329 ----- 329  
Db 3777 CGGAAGTACAGCATGCGGAGACTGCTGCAGAAACGAGACTGTGAGCCGCTGACACT 3836  
QY 329 ----- 329  
Db 3837 AGCGAGCGATGCCAACACAGCGCAGATGCGGATCTCTAAAGAGACGAGAGCTGAGAAAG 3896  
QY 329 ----- 329  
Db 3897 GTGAAGTGTCTTGATCTGGCGCTTTTGGCAGAGTCTACAAAGGCACTGATCCCTGAT 3956  
QY 329 ----- 329  
Db 3957 GGGGAGATGTGAAAATTTCCAGTGGCCATCAAAAGTTTGAGGGAAAACACATCCCCCAA 4016  
QY 329 ----- 329  
Db 4017 GCCAACAAGAAATCTTAGACGAGCATACGTGATGGTGTGTGGCTCCCATATATGC 4076  
QY 329 ----- 329  
Db 4077 TCCCGCTTCTGGGATCTGCTGACATCCAGGTGCAGCTGTGACACAGCTTATGCC 4136  
QY 329 ----- 329  
Db 4137 TATGGCTGCTTTAGACCATGTCCGGGAAAACCGGACGCTGGGCTCCAGACACTG 4196  
QY 329 ----- 329  
Db 4197 CTGAAGTGTATGACAGATTCGAAGGGATGAGCTACCTGAGAGATGTGGGCTGTA 4256  
QY 329 ----- 329  
Db 4257 CACAGGACTTGGCGCTCGGAACGTCTGTCAAGAGTCCCAACCATGTCAAAATTACA 4316  
QY 329 ----- 329  
Db 4317 GACTTCGGGCTGGCTCGCTGCTGACATTGACGAGACAGATACCATCAGATGGGGGC 4376  
QY 329 ----- 329  
Db 4377 AAGTTCACATCAAGTATGATGGCGCTGAGTCCATTCTCCGCGGCTTCAACCCACAG 4436  
QY 329 ----- 329  
Db 4437 AGTATGTGTGAGTTATGCTGTGACTGTGTGGAGCTGATGACTTTGGGGCCAAACT 4496  
QY 329 ----- 329  
Db 4497 TACGATGGATCCAGCCCGGAGATCCCTGACCTGTGGAAAAGGGGAGGGGCTGCC 4556  
QY 329 ----- 329  
Db 4557 CAGCCCCCATGTGCACATGATGTCTACATGATGTGTAATGTTGATGATTGAC 4616  
QY 329 ----- 329  
Db 4617 TCTGAATGTGGCCAAAGATTCGGGAGTTGTGTCTGAATTCTCCGCAATGCCAGGAGAC 4676  
QY 329 ----- 329  
Db 4677 CCCCAGCCTTTGTGTATCCAGAAATGAGACTTGGGCCAAGCATCCCTTGACAGC 4736  
QY 329 ----- 329  
Db 4737 ACCTTACCGCTCACTGCTGAGAGCATGACATGGGGAGCTGTGATGCTGAGAG 4796  
QY 330 ----- 334  
Db 4797 TATCTGTACCCAGCAGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATG 4856

```

Oy 335 VALHSHARGHISARGSerSerSerThrArgSerGlyGlyAspLeuThrLeuGly 354
Db 4857 GTCCACACAGGACCGACGACATCTACACAGAGTGGCGGGGAGCTGACCTAGG 4916
Oy 355 LeuGIuProSerGIuGIuAlaProArSerProLeuAlaProSerGIuGIuAlaGly 374
Db 4917 CTGGAGCCCTCTGAGAGAGAGGCCCCCAAGCTCTCCACTGGCAGCCTCCGAA 4976
Oy 375 SerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuPro 394
Db 4977 TCCGATGATTATGATGATGACCTGGGAATGGGGACCAAGGGGCTGCAGAAAGCTCC 5036
Oy 395 ThrHisAspProSerProLeuGlnArgTyrSerGIuAspProThrValProLeuProSer 414
Db 5037 ACACATGACCCCAAGCCCTCTACAGCGGTACAGTGAAGGACCCCAAGTACCTGCTCC 5096
Oy 415 GluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAsnGln 434
Db 5097 GAGACTGATGGCTACGTTGCCCTGACCTGACGACCCCAAGCTGAAATATGTGAACAG 5156
Oy 435 ProAspValArgProGlnProProSerProArGluGlyProLeuProAlaAlaArgPro 454
Db 5157 CCAGATGTTCCGGCCCAAGCCCTCTCCGCGGAGAGGGCCCTGCTGCTGCGCCGACCT 5216
Oy 455 AlaGlyAlaThrLeuGIuArgAlaAlaThrLeuSerProGlyValAsnGlyValAlaGly 474
Db 5217 GCTGATGCCACTCTGGAAGGACCAAGACTCTCTCCCAAGGAAAGGGGTCTCTCAAA 5276
Oy 475 AspValPheAlaPheGlyGlyValAlaValGluAsnProGluTyrLeuThrProGlnGly 494
Db 5277 GACGTTTGGCTTGGGGGGTGGCGTGAAGAACCCGAGTACTTGACACCCCAAGGAGGA 5336
Oy 495 AlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 514
Db 5337 GCTGCCCTCAGCCCAAGCCCTCTCTGCTTCAAGCCCAAGCTTCAAGCTCTATATAC 5396
Oy 515 TrpAspGlnAspProProGluArgGlyAlaProProSerThrPheGlyGlyThrProThr 534
Db 5397 TGGGACCAAGGACCCCAAGAGCGGGGGCTCCACCAAGCTTCAAGGAGCACTTACG 5456
Oy 535 AlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSerPro 554
Db 5457 GCAGAGAACCCAGAGTACTGGGTCTGAGGTGCCA-----GTGTGAACGAGAAAGGCCA 5510
Oy 555 SerProSerThrGlnProTyrGluHisValAsnAlaAlaGlnGluAlaArgArgLeu 574
Db 5511 AGTCCGCGAG---AAGCCCTGATGTCTCTCAGG-----GAGCAGGAGGAGCGGCTCTG 5561
Oy 575 AsnLeuSerArgAspThr 580
Db 5562 AGCTATTCAGAGAGTACT 5579

```

```

RESULT 9
US-09-441-411-5
: Sequence 5, Application US/09441411
: Publication No. US2003008342A1
: GENERAL INFORMATION:
: APPLICANT: Scholler, Nathalie B.
: APPLICANT: Dais, Mary L.
: APPLICANT: Helstrom, Ingegerd
: APPLICANT: Helstrom, Karl Erik
: TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
: FILE REFERENCE: 730033.409
: CURRENT APPLICATION NUMBER: US/09/441.411
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 4473
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-441-411-5

```

```

Alignment Scores:
Pred. No.: 3,44e-200 Length: 4473
Score: 2418.00 Matches: 523
Percent Similarity: 41.06% Conservative: 5
Best Local Similarity: 40.67% Mismatches: 22
Query Match: 64.67% Indels: 736
DB: 9 Gaps: 4

US-09-821-883-2 (1-690) x US-09-441-411-5 (1-4473)
Oy 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 118 CGCGCTCTCCCAAGCCGGGTCCAGCCGGAGCCATGGGCGCCGAGCCGACGTGACACCTG 177
Oy 18 PheLeuPheLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 37
Db 178 GAGCTGGGGGGCTTGTGGCGGTG-----GGGCTCTCTCTGCGCCCTTGTGGCCCCC 228
Oy 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 229 GGAGCGCGGACACCCCAAGTGTGCACCGCACAGACATGAAGCTCGGCTCTCCCTCCAGT 288
Oy 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyGlnValValGlnGly 77
Db 289 CCCGAGACCCACCTGACATGCTCCGCCACTTACACAGGCTGCGAGTGTGACAGGA 348
Oy 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 349 AACCTGGAACCTACCTACTCTGCCCAACATGCCAGCTGCTCTCTCTCCAGATATCCAG 408
Oy 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
Db 409 GAGGTGACGGGCTACGTCTATGCTCACAACCAAGAGAGGAGGTCCACTGCACAGG 468
Oy 118 LeuArgIleValAlaGlyGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
Db 469 CTGGGATGTGGCAGAGGACACCACTCTTTGAGGACAACTATGCCCTGCGCTGCTACAC 528
Oy 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 529 AATGAGACCCCGTGAACAAATACCAACCCCTGTACAGAGGGCTCCCGCAGAGGCTCGG 588
Oy 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuGlyGlyValLeuIleGlnArgAsn 177
Db 589 GAGCTGACGCTTCAAGCTCAGAGCTCAAGATCTTGAAGAGGGGTGTATCACCAGGAC 648
Oy 178 ProGlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGln 197
Db 649 CCCAGCTCTGCTACACAGGACACGATTTTGTGAAGGACATCTTCCACAAACACACAG 708
Oy 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPromet 217
Db 709 CTGGCTCTCACACTGATGATGACACCAACCGCTCTGGGCTCCACCCCTGTTCTCCGATG 768
Oy 218 CysLysGlySerArgCysTyrGlyGluSerSerGlnAspCysGlnSerLeuThrArgThr 237
Db 769 TGTAAAGGCTCCCTGCTGCTGGGAGAGGTCTGAAGATTTCTCAAGCTCAAGCTCAAGGCT 828
Oy 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGln 257
Db 829 GTCTGTGGCGGTGGCTGCGCCGCTGCAAGGGGACAGTCCACTGACTGCTGCAGAGAG 888
Oy 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 889 CAGTGTGTGCGGCTGACACGGGCCCCCAAGACACTCTGCTGCGGCTGCTCCACTTTC 948
Oy 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValIleTyrAsnThrAspThr 297
Db 949 AACCAAGTGGCATCTGTGAGCTCACTGCGCCAGCCCTGCTCAGCTTCAACACAGACAG 1008
Oy 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
Db 1009 TTTGAGTCCATGCCCATTCGAGAGGGCGGTATACATTCGCGGCGGCTGCTGACTGCC 1068

```



QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329  
Db 1069 TETCCCTACACTACTCTTCTACGAGCGTGGATCTCTGCACCCCTGCTGCCCCCTGCAC 1128  
QY 329 ----- 329  
Db 1129 AACCAAGAGTGCACAGAGAGTGAACACAGCGGTGTGAGAAATGCAGCAAGCCCTGT 1188  
QY 329 ----- 329  
Db 1189 GCCCAGTGTGCTATGTGCTGTGGCATGAGACTTGCAGAGGTGAGGGCAGTTACCACT 1248  
QY 329 ----- 329  
Db 1249 GCCCATATCCAGAGAGTTTCTGCTGCTGCAGAAAGATCTTTGGAGCCTGGCATTTCTGCCG 1308  
QY 329 ----- 329  
Db 1309 GAGAGCTTGTATGGGAGCCAGCCTCCACACTGCCCGCTCCAGCCAGAGAGCTCCAA 1368  
QY 329 ----- 329  
Db 1369 GTGTTGAGACTCTGAGAGAGATCACAGGTTACTATACATCTCAGCATGGCCGGAGAC 1428  
QY 329 ----- 329  
Db 1429 CTGCCTGACCTCAGCGCTTCTCCAGAACTGCAGATATCCGGGGAGCAATTTCTGCACAT 1488  
QY 329 ----- 329  
Db 1489 GCGGCTTACTCGCTGACCCCTGCAAGGGGTGGGCAATCAGCTGCTGGGCTGCGCTACTG 1548  
QY 329 ----- 329  
Db 1549 AAGGAATGGGAGTGGAGTGGCCCTCATCCACCATACACCCACTCTGCTTCGTGCAC 1608  
QY 329 ----- 329  
Db 1609 ACGGTGCCCTGGGACACAGCTTTTGGAAACCCGACCAAGCTCTGTCTCAGACTGCCAAC 1668  
QY 329 ----- 329  
Db 1669 CGGCCAGAGAGAGTGTGTGGGCGAGGGCTGGCCTGCCACAGCTGTGCGCCGAGGG 1728  
QY 329 ----- 329  
Db 1729 CACTGCTGGGTCCAGGGCCACCCAGTGTGTCACTGCAGCCAGTTCTTGGGGCCAG 1788  
QY 329 ----- 329  
Db 1789 GAGTGTGGAGGAATGCCGAGTACTGTCAAGGGGCTCCAGAGGAGTATGTGAATGCCAGG 1848  
QY 329 ----- 329  
Db 1849 CACTGTTTGCCTGACCCCTGAGTGTCAAGCCCAAGATGGCTCAGTGACTGTTTGA 1908  
QY 329 ----- 329  
Db 1909 CCGGAGGCTGACCACTGTGTGGCTGTGCCCACTATAGAGACCTCCCTTCTGCGTGGCC 1968  
QY 329 ----- 329  
Db 1969 CCCTGCCCCAGGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2028  
QY 329 ----- 329  
Db 2029 GAGAGGGGCGCATGCCAGCCTTGCCCCATCACTGCACCCACTCTGTGTGACCTGAT 2088  
QY 329 ----- 329  
Db 2089 GACAAAGGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTG 2148

QY 329 ----- 329  
Db 2149 GTTGGCATTTCTGCTGTGTGCTCTTGGGGGTGTGTTGGATCCTCATCAAGGAGCG 2208  
QY 329 ----- 329  
Db 2209 CAGCAGAAAGATCCGGAAGTACAGATGCGGAGACTGTCTGAGAAACGAGCTGTGTAG 2268  
QY 329 ----- 329  
Db 2269 CCGCTGACACTTACGGGAGCCATGCCCAACAGGGCGAGATCGGATCCTGAAAGAGACG 2328  
QY 329 ----- 329  
Db 2329 GAGCTGAGGAAGTGAAGTGTGATGTGGGCTTTTGGCAGTCTACAAAGGCAATC 2388  
QY 329 ----- 329  
Db 2389 TGGATCCCTGATGGGAGAAATGAAATTCAGTGGCCATCAAGTGTGAGGAAAC 2448  
QY 329 ----- 329  
Db 2449 ACATCCCCAAAGCCACAAAGAAATCTTAGACGAAGCATACGTATGGCTGTGTGGC 2508  
QY 329 ----- 329  
Db 2509 TCCCATATGTCTCCCGCTTGTGGCATCTGCTGACATCCAGGCTGCAGCTGGTGACA 2568  
QY 329 ----- 329  
Db 2569 CAGCTTATCCCTATGGCTGCTTAGACATGTCCGGGAAACCGCGAGCGCTGGGC 2628  
QY 329 ----- 329  
Db 2629 TCCAGAGACTGCTGAATGTGTATGATGCAATTTGCCAAGGGATGACTACCTGGAGAT 2688  
QY 329 ----- 329  
Db 2689 GTGGGCTGTACACAGGAGCTTGGCCGCTCGGAACGTGTGTCMAAGATCCCAACAT 2748  
QY 329 ----- 329  
Db 2749 GTCAAAATTAACAGACTTCGGGCTGGCTCGCTGTGACATTGACGAGACAGATACAT 2808  
QY 329 ----- 329  
Db 2809 GCAGATGGGGCAAGTGCCTCATCAAGTGAATGGCGCTGGAATCATTTCCGCCGGCGG 2868  
QY 329 ----- 329  
Db 2869 TTCACCCACAGAGTATGTGTGAGTTATGTGTGACTGTGTGGAGGCTGATGACTTTT 2928  
QY 329 ----- 329  
Db 2929 GGGGCCAAACTTACAGTGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGG 2988  
QY 329 ----- 329  
Db 2989 GAGCGGCTCCCGAGCCCGCCCATCTGCACCAATTGATGTCTACATGATCATGTCAAAATGT 3048  
QY 329 ----- 329  
Db 3049 TGGATGATTGACTGTGAATGTGGGCCAAGATTCCGGGATTTGTTGTCTGAATTTCCCGC 3108  
QY 329 ----- 329  
Db 3109 ATGGCAGGGAGCCCGAGCGCTTGTGTCAATCAGAAATGAGAGACTTGGGCCAGCCAGT 3168  
QY 329 ----- 329  
Db 3169 CCCTTGGACAGCACTTCTACCGCTACTGTCTGAGAGACAGATGACATGAGGGAGCTGTG 3228  
QY 330 -----Gly 330



QY 329 ----- 329  
Db 946 CCCCTGCACAACAAGAGGTGACAGAGAGATGGAACACAGCGGTGTGAGAAATGACAC 1005  
QY 329 ----- 329  
Db 1006 AAGCCCTGTGCCAGAGTGTCTATGTCTGTGGCATGAGACACTTGCAGAGAGTGAGGGCA 1065  
QY 329 ----- 329  
Db 1066 GTTACAGATGCCAATATCCAGAGAGTTGTGCTGTGCAGAAAGATCTTTGGAGCCTGSCA 1125  
QY 329 ----- 329  
Db 1126 TTTCGCCGAGAGAGCTTTGATGGGAGCCACCTCCAACTGTGCCCGCTCCAGCCAGAG 1185  
QY 329 ----- 329  
Db 1186 CAGCTCCAAATGTTTGAAGCTGTGAGAGATCACAGGTTACTATACATCTCAGCATGG 1245  
QY 329 ----- 329  
Db 1246 CCGGACAGCTTCCCTGACCTCAGCGTCTTCCAGAACTGCAAGTATCCGGGGAGCAATT 1305  
QY 329 ----- 329  
Db 1306 CTGCACATGGCGCTACTCGCTGACCTGCAAGGCTGGGCAATCAGCTGGGCTGGGCTG 1365  
QY 329 ----- 329  
Db 1366 CGCTACTGAGGAACTGGGCAATGAGACTGGCCCTCATCACATACACCCACCTCTGC 1425  
QY 329 ----- 329  
Db 1426 TTCTGCACAGGAGTCCCTGGAGACAGCTCTTTGGGAACCGGACCAAGCTCTGCTCAC 1485  
QY 329 ----- 329  
Db 1486 ACTGCACACGGGCGAGAGAGAGATGTGTGGGCGAGGGCTGGGCTGCACACAGCTGTC 1545  
QY 329 ----- 329  
Db 1546 GCCGAGGCGACTGTGGGTCCAGGGCCACCCAGTGTGTCACTGACGACAGTTCCTT 1605  
QY 329 ----- 329  
Db 1606 CGGGGCCAGAGTGGGTGAGAGAAATGCCAGTACTGTCAGAGGGCTCCAGAGGATATGTG 1665  
QY 329 ----- 329  
Db 1666 AATGCGACGACACTGTTTCCGTGCCACACCTGAGTGTCAAGCCCAAGATGGCTCAGTAC 1725  
QY 329 ----- 329  
Db 1726 TGTTTGACCGAGAGCTGACCAAGTGTGTGCTGTGCCACTATAGAGACCTCCCTTC 1785  
QY 329 ----- 329  
Db 1786 TGGGTGGCCGCTGCCCAAGGCTGTGAACCTGACTCTCTACATGCCATCTGGAAAG 1845  
QY 329 ----- 329  
Db 1846 TTTCAGATGAGAGAGGCGCATGCCAGCTTGGCCCATCACTGACCCACACTCTGTGTG 1905  
QY 329 ----- 329  
Db 1906 GACCTGATGACAAAGGCTGCCCGCGAGAGAGAGCCACCTGTGACATCATCTGTC 1965  
QY 329 ----- 329  
Db 1966 TCTGCGGTGGTTGGCATTTCTGTGTGCTCTTGGGGGTGGTCTTTGGGATCCTCATC 2025  
QY 329 ----- 329

Db 2026 AAGGACGGCAGAGAAAGATCCGGAAGTACAGATGCCGAGACTGTCCAGAAAAAGAG 2085  
QY 329 ----- 329  
Db 2086 CTGTGAGCCGCTGACACACTAGCGAGCGATGCCAACAGGGCGAGATCCGATCCTG 2145  
QY 329 ----- 329  
Db 2146 AAAGAGCGAGCTGAGGAAGTGAAGTGTGTGATCTGGCGCTTTTGGCACACTTAC 2205  
QY 329 ----- 329  
Db 2206 AAGGCACTGTGATCCCTGATGGGAGAAATGAAAAATTCAGTGGCATCAAAATGTGTG 2265  
QY 329 ----- 329  
Db 2286 AAGGAAACACATCCCGCCAAAGCCAAAGAAATCTTAGACAAAGCATACGTATGGCT 2325  
QY 329 ----- 329  
Db 2326 GGTGTGGGCTCCCATATATGTCTCCCGCTCTGGGCATCTGCTGACATCCAGGCTGCG 2385  
QY 329 ----- 329  
Db 2386 CTGTTGACACAGCTTATGCTCCTATGCTGCTCTTAGACATGTCCGGAAACCGCGGA 2445  
QY 329 ----- 329  
Db 2446 CGCTGGGCTCCAGAGACCTGCTGAATGTGTATGAGAAATGCCAAGGGAATGAGTAC 2505  
QY 329 ----- 329  
Db 2506 CTGAGAGATGTGGGCTGTACACAGGACCTTGGCCGCTCGGAACGTCTGTCAAGAT 2565  
QY 329 ----- 329  
Db 2566 CCCAACATGTCAAAATATACAGACTTGGGCTGGCTCGCTGTGACATTGACGAGACA 2625  
QY 329 ----- 329  
Db 2626 GAGTACATGAGATGGGGGAGAGGTGCCATCAAGTGAATGGCGGTGATCTCTC 2685  
QY 329 ----- 329  
Db 2686 CGCGGGGCTTACCCACACAGATGATGTGAGATTATGTGTACTGTGTGGAGCTG 2745  
QY 329 ----- 329  
Db 2746 ATGACTTTTGGGGCCAAACTTACGATGGATCCAGCCGAGAGATCCATGACCTGCTG 2805  
QY 329 ----- 329  
Db 2806 GAAAAAGGGAGCGGCTCCCGAGCCCGCATCTGCACCATTTGATGTCTACATCATC 2865  
QY 329 ----- 329  
Db 2866 GTCAATGTGGATGATTTGACTCTGAATGTGCGGCCAAGATTCGGGAGTTGTGTCTGAA 2925  
QY 329 ----- 329  
Db 2926 TTCTCCGATAGGACGAGGACCCCGAGCGCTTGTGTATTCAGAAATGAGAGACTTGGGC 2985  
QY 329 ----- 329  
Db 2986 CCAGCACTCCCTTGGACAGACACTTCTACCGCTCTACGCTGTGAGAGAGATGACATGGGG 3045  
QY 329 ----- 329  
Db 3046 GACCTGTGATGCTGAGGAGATATCTGTACCCAGAGGGCTTCTGTGTCCAGACCT 3105  
QY 330 ----- 347  
-----GlyAlaGlyGlyMetValHisHisArgHisArgSerSerThrArgSerGly  
|||||

Db 3106 GCCCGGGGGCGGCGATGTCACACGACGACCGACGCTCATCTACAGAGTGGC 3165  
Qy 348 G1yG1yAspLeuThrLeuG1yLeuG1uProSerG1uG1uAlaProArgSerProLeu 367  
Db 3166 GGTGGGACCTGACACTAGGGGCTGGAGCCCTCTGTAAGAGAGGGCCCGAGTCTCCACTG 3225  
Qy 368 AlaProSerG1uG1yAlaG1ySerAspValPheAspG1yAspLeuG1yMetG1yAlaAla 387  
Db 3226 GCACCCCTCCAGAGGGGCTGCTCCGATGTATTGATGTGATGCTGGAATGGGGGAGCC 3285  
Qy 388 LysG1yLeuG1uInserLeuProThrHisAspProSerProLeuG1uArgTySerG1uAsp 407  
Db 3286 AAGGGGCTGCAAGCTCTCCACACATGACCCAGCCCTCTACAGCGGTAAGTGAAGGAC 3345  
Qy 408 ProThrValProLeuProSerG1uThrAspG1yTyValAlaProLeuThrCysSerPro 427  
Db 3346 CCCACATACCCCTGCTGAGACTGATGGTGGCTGCGCCCGCCCTGACCTGAGCCCGC 3405  
Qy 428 GlnProG1uTyValAlaGlnProAspValArgProGlnProProSerProArgG1uG1y 447  
Db 3406 CACCTGAAATATGTAACACAGCAGATGTTCGGCCCGACCCCTTCGCCCGAGAGGCG 3465  
Qy 448 ProLeuProAlaAlaArgProAlaG1yAlaThrLeuG1uArgAlaLysThrLeuSerPro 467  
Db 3466 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3525  
Qy 468 G1yLysAsnG1yValValLysAspValPheAlaPheG1yAlaValG1uAsnProG1u 487  
Db 3526 GGGAAAGATGGGGCTGCAAGAGCTTTTGGCTTTGGGGGTGGCGGAGAACCCCGAG 3585  
Qy 488 TyLeuThrProG1uG1yG1yAlaAlaProGlnProHisProProAlaPheSerPro 507  
Db 3586 TACTGTACACCCAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3645  
Qy 508 AlaPheAspAsnLeuTy 527  
Db 3646 GCCTTGACAACTCTTACTGTGAGACAGACAGACACCAAGCGGGGCTCCACCCAGC 3705  
Qy 528 ThrPheLysG1yThrProThrAlaG1uAsnProG1uTyLeuG1yLeuAspValPro 546  
Db 3706 ACCTTCAAAGGAGACCTACGCGAGAGAACCCAGAGTACCTGGGTGAGAGTCCCA 3762

RESULT 11  
US-09-811-115-2  
Sequence 2, Application US/09811115  
Patent No. US20020035736A1  
GENERAL INFORMATION:  
APPLICANT: Erickson, Sharon  
APPLICANT: Schwall, Ralph  
APPLICANT: King, Kathleen  
TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
FILE REFERENCE: GENENT. 034A  
CURRENT APPLICATION NUMBER: US/09/811,115  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/189,844  
PRIOR FILING DATE: 2000-03-16  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 3768  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-811-115-2

## Alignment Scores:

Pred. No.: 1 259-199 Length: 3768  
Score: 2410.50 Matches: 510  
Percent Similarity: 41.16% Conservative: 0  
Best Local Similarity: 41.16% Mismatches: 2  
Query Match: 64.47% Indels: 727  
DB: 10 Gaps: 1

US-09-821-883-2 (1-690) x US-09-811-115-2 (1-3768)

Qy 35 LeuAlaArgG1yAlaAlaSerThrGlnValCysThrG1yThrAspMetLysLeuArgLeu 54  
Db 46 TTGGCCCCCGGAGCGGAGAGACCCAGTGTGACCGCGACAGACATGAAAGCTCGGCTC 105  
Qy 55 ProAlaSerProG1uThrHisLeuAspMetLeuArgHisLeuTyGlnG1yCysGlnVal 74  
Db 106 CTTGCCAGTCCCGAGACCCACCTTGACATGCTCCGCCACCTCTACACAGGGCTGCGAGGTG 165  
Qy 75 ValGlnG1yAsnLeuG1uLeuThrTyLeuProThrAsnAlaSerLeuSerPheLeuGln 94  
Db 166 GTGCAGGGAACCTCGAAGCTACCTACCTGCCCCCAATGCGACCTGCTCTCTGAG 225  
Qy 95 AspIleGlnG1uValGlnG1yTyValLeuIleAlaHisAsnGlnValArgGlnValPro 114  
Db 226 GATATCCAGGAGGTGACGAGGCTACGTGCTGATGCTGCATCCATCAACCAAGAGAGAGTCCCA 285  
Qy 115 LeuGlnArgLeuArgIleValArgG1yThrGlnLeuPheG1uAspAsnTyValAlaAla 134  
Db 286 CTGCAGAGGCTGCGGATGTGCGAGGCCACCAAGCTTTTGAAGCAACTATGCGCTGGCC 345  
Qy 135 ValLeuAspAsnG1yAspProLeuAspAsnThrThrProValThrG1yAlaSerProG1y 154  
Db 346 GTGCTAACAAATGAGACCCGCTGTAACAAATACCAACCTGTACAGAGGCGCTCCAGCA 405  
Qy 155 G1yLeuArgG1uLeuGlnLeuArgSerLeuThrG1uLeuLysG1yG1yValLeuIle 174  
Db 406 GGCCTGCGGAGGCTGACGCTTCGAAAGCTCACAAGATCTTGAAGAGAGGGGTGTATC 465  
Qy 175 GlnArgAsnProGlnLeuCysTyGlnAspThrIleLeuThrLysAspIlePheHisLys 194  
Db 466 CAGCGGAACCCCGACGCTGTGTACAGACACGATTTTGTGAAGACATCTTCACAG 525  
Qy 195 AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCys 214  
Db 526 AACAGACAGGTGGCTGCACACTGATGACACCAACCGCTGTGGGCTGCACACCTGT 585  
Qy 215 SerProMetCysLysG1ySerArgCysTrpG1yLysSerG1uAspCysGlnSerLeu 234  
Db 586 TCTCCGATGTGTAAGGCTCCCGCTGCTGGGAGAGAGTTCTGAGATTCAGAGCTG 645  
Qy 235 ThrArgThrValCysAlaG1yG1yCysAlaArgCysLysG1yProLeuProThrAspCys 254  
Db 646 ACGCGACTGTCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705  
Qy 255 CysHisGlnG1uCysAlaAlaG1yCysThrG1yProLysHisSerAspCysLeuAlaCys 274  
Db 706 TGCATAGAGCAGTGTGCTGCGGCTGCGAGCGGCCCAAGACCTGTGACTGCTGCTGCTG 765  
Qy 275 LeuHisPheAsnHisSerG1yIleCysGlnLeuHisCysProAlaLeuValThrTyAsn 294  
Db 766 CTCACACTCAACACACAGCTGACATCTGTGACTGTGACAGCCCAAGCTGTGACTTACAC 825  
Qy 295 ThrAspThrPheG1uSerMetProAsnProGlnG1yArgTyThrPheG1yAlaSerCys 314  
Db 826 ACAGACAGCTTGTGATCCATAGCCCAATCCGAGGGCGGATATCATTCGCGGCGCAGCTGT 885  
Qy 315 ValThrAlaCysProTy 329  
Db 886 GTAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945  
Qy 329 ----- 329  
Db 946 CCCGTGCAACCAAGAGGTGACAGCAGAGATGAAACACAGCGGTGTGAGAAATGCGAGC 1005  
Qy 329 ----- 329  
Db 1006 AAGCCCTGTGCGGAGTGTCTATGTCTGTGGCATGAGACACTTGGCAGAGGTGAGGCA 1065  
Qy 329 ----- 329  
Db 1066 GTTACACAGTGCCAATATCCAGAGATTGCTGCTGCGCATGAGAAAGATCTTTGGAGACCTGGCA 1125

QY 329 ----- 329  
Db 1126 TTTTCGCGGAGAGCTTTGATGGGAGCCAGCTCCAACTGCCCCGCTCCAGCCAGAG 1185  
QY 329 ----- 329  
Db 1186 CAGCTCAAGTGTGAGACTGTGAGAGATCACAGTTACTATACATCTCAGCATGG 1245  
QY 329 ----- 329  
Db 1246 CCGGACAGCTGCTGACCTCAGCGTCTTCAGAACTGCAAGTAATCCGGGGAGCAATT 1305  
QY 329 ----- 329  
Db 1306 CTGCACATGGGGCTACTCGCTGACCTGCAAGGGCTGGCATGAGTGGGCTG 1365  
QY 329 ----- 329  
Db 1366 CGCTCACTGAGGGAAGTGGGAGTGGAGTGGCTTCATCCACCATTAACACACCTCTGC 1425  
QY 329 ----- 329  
Db 1426 TTCTGACACAGGTGCTCCCTGGAGCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCAC 1485  
QY 329 ----- 329  
Db 1486 ACTGCCAACCGGCGCAGAGAGCAGTGTGTGGGGGAGGGGCTGGCTGCCACCACTGTGC 1545  
QY 329 ----- 329  
Db 1546 GCGCGAGGCACTGCTGGGGTCCAGGGCCACCCAGTGTGTCACATGACAGATTCTT 1605  
QY 329 ----- 329  
Db 1606 CCGGCGCAGAGTGGCTGGAGGAATCCGAGTACTGACAGGGCTCCCCAGGAGTATGTG 1665  
QY 329 ----- 329  
Db 1666 AATGCCAGGCACTTTTGGCGGTGCCACCTGATGTCAAGCCCAAGATGGCTCAGTGAAC 1725  
QY 329 ----- 329  
Db 1726 TGTTTTGGACCGAGGCTGACAGTGTGTGGCTGTGCCACTAATAAGAGCCCTCCCTTC 1785  
QY 329 ----- 329  
Db 1786 TGCCTGCCCCGCTGCCCGCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAA 1845  
QY 329 ----- 329  
Db 1846 TTTCCAGATGAGAGGGGCGCATGCGAGCCTTGCCTCCATCAACTGCACCCACTCTGTGTG 1905  
QY 329 ----- 329  
Db 1906 GACCTGATGACAGAGGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACATCCATCTGC 1965  
QY 329 ----- 329  
Db 1966 TCTGGGCTGTTGGCATTTCTGCTGTGTGTCTGGGGGTGTCTTTGGGATCCTCATTC 2025  
QY 329 ----- 329  
Db 2026 AAGCGAGCGCAGCAGAAAGATCCGGAAGTACAGATGCGGAGACTGCTGAGGAAGCGGAG 2085  
QY 329 ----- 329  
Db 2086 CTGTGAGCCGCTGACACCTAGCGAGCGATGCCCAACAGCGCAGATGCGGATCTCTG 2145  
QY 329 ----- 329  
Db 2146 AAAGAGAGGAGCTGAGGAAGGTGAAGGTCTTGATCTGGCGCTTTTGGGACAGCTTAC 2205  
QY 329 ----- 329

Db 2206 AAGGCAATCTGATCCCTGATGGGAGATGTGAAATTCAGTGGCCATCAAACTGTG 2265  
QY 329 ----- 329  
Db 2266 AAGGAAACACATCCCGCAAAACCAAGAAATTCAGAGAGATACGTGATGCT 2325  
QY 329 ----- 329  
Db 2326 GGTGTGGGCTCCCATATGTCTCCGCGCTTCTGGGCATCTGCCTGACATCCAGGTCAG 2385  
QY 329 ----- 329  
Db 2386 CTGTGACACAGCTTATGCGCTATGCTGCTCTTAGACCATGTCCGGGAAAAACCGCGGA 2445  
QY 329 ----- 329  
Db 2446 CGCTGGGCTCCAGAGACCTGCTGAACGTGTGTATGCAAGATTGCCAAGGGATGAGCTAC 2505  
QY 329 ----- 329  
Db 2506 CTGAGAGATGTGGGCTGTACACAGGACTTGGCCGCTCGGACAGTGTCTCAAGAT 2565  
QY 329 ----- 329  
Db 2566 CCCAACATGTCAAAATTCAGACTTCGGGCTGGGCTCGGCTCTGACATTGACGAGACA 2625  
QY 329 ----- 329  
Db 2626 GAGTACATGAGATGGGGGCAAGTGCCCATCAAGTGAATGGCGCTGATCCATTCTC 2685  
QY 329 ----- 329  
Db 2686 CCGCGGCGGTTACCCACACAGTATGTGTGAGTTATGTGTACTGTGGGAGCTG 2745  
QY 329 ----- 329  
Db 2746 ATGACTTTTGGGGCCAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACTGTG 2805  
QY 329 ----- 329  
Db 2806 GAAAGGGGAGCGGCTGCCAGCCCGCATCTGCACATTGATGTATCATGATCATG 2865  
QY 329 ----- 329  
Db 2866 GTCAATGTGATGATTTACTCTGATGTGCGCCAAAGATTCCGGAGTGTGTCTGA 2925  
QY 329 ----- 329  
Db 2926 TTCTCCGATGCGCAGGAGCCCGAGCGCTTTGTGTATCCAGAAATGAGCATTTGGGC 2985  
QY 329 ----- 329  
Db 2986 CCAGCCAGTCCCTGGAGACACCTTCTACCGCTCACTGTGAGAGAGCATGACATGGG 3045  
QY 329 ----- 329  
Db 3046 GACCTGTGATGCTGAGAGATCTGTGATCCCGACGAGGCTTCTTGTCTCAGAACCT 3105  
QY 330 ----- 347  
Db 3106 GCGCGGCGGCTGGGGGATGCTCCACACAGCAGCCGAGCTCATCTACAGAGATGTC 3165  
QY 348 G1yG1yAspLeuTh1eug1yLeuG1uProSerG1uG1uAla1PArArgSerPro1eu 367  
Db 3166 GGTGGGAGCTTACACTAGGGCTTGAGCCCTTGAAGAGAGAGCCCGCGTCTCCACAG 3225  
QY 368 AlAProSerG1uG1yAlaG1ySerAsp1yPheAspG1yAspLeuG1yMetG1yAla1a 387  
Db 3226 GCACCTCCGAAGGGGCTGGCTCCATGTATTTGATGTGACACTGGGAATGGGGCAGCC 3285  
QY 388 LysG1yLeuG1nSerLeuProThr1yAspProSerProLeuG1nArgTyrSerG1uAsp 407  
|||||



QY 329 ----- 329  
Db 1006 AAGCCCTGTGCCGAGTGTGCTATGCTCTGGGCATGAGACACTTGGAGAGGTGAGGGCA 1065  
QY 329 ----- 329  
Db 1066 GTTACAGTGCATATATCCAGAGAGTCTTGCTGCTGCAAGAAGATCTTTGGAGCCTGGCA 1125  
QY 329 ----- 329  
Db 1126 TTTCTGCCGAGAGCTTTGATGGGACCAGCCTCCAACTGCCCCGCTCCAGCCAGAG 1185  
QY 329 ----- 329  
Db 1186 CAGCTCCAGTGTTTAGAGCTCTGGAAGATCAGAGTTACCTATACATCTCAGCATGG 1245  
QY 329 ----- 329  
Db 1246 CCGGACAGCTGTCTGACCTCAGCGTCTCCAGAACCTGCAAGTATCCGGGAGCAAT 1305  
QY 329 ----- 329  
Db 1306 CTGACATAGGCGCTACTGCTGACCTGCAAGGCTGGGCATCAGCTGGGCTG 1365  
QY 329 ----- 329  
Db 1366 CGCTCACTGAGGGAAGTGGGAGTGAAGTGGCCCTCATCCACATTAACACCCACTCTGC 1425  
QY 329 ----- 329  
Db 1426 TTGCTGACACAGGTGCCCTGGGACAGCTCTTTCGGAACCGCACCAAGTCTGCTCCAC 1485  
QY 329 ----- 329  
Db 1486 ACTGCCAACCGGCCAGAGAGAGAGTGTGTGGGCGAGGGCTTGCCCTGACCACACTGTGC 1545  
QY 329 ----- 329  
Db 1546 GCCCGAGGACACTGCTGGGCTCAGGGCCACCAAGTGTGTCACTGAGCCAGTTCCTT 1605  
QY 329 ----- 329  
Db 1606 CGGGCCAGAGTGTGCTGAGAGAAATGCCAGTACTGAGGGGCTCCCAAGGAGTATGTG 1665  
QY 329 ----- 329  
Db 1666 AATGCCAGGCACTGTTGCCGTGCCACCCTGAGTGTGACGCCCCAGAAATGGCTCAGTGACC 1725  
QY 329 ----- 329  
Db 1726 TGTTTGGACCGAGGCTGACAGTGTGTGGCTGTGCCACTATAAGAGACCTCCCTTC 1785  
QY 329 ----- 329  
Db 1786 TGCCTGCCCGCTGCCCAAGCGGTGTGAACCTGACTCTCTCTACATGCCCATCTGGAAG 1845  
QY 329 ----- 329  
Db 1846 TTTCCAGATGAGAGGGGCGCATGCGACCTTGCCCATCAACTGCACCCACTCTGTGTG 1905  
QY 329 ----- 329  
Db 1906 GACCTGATGACAAAGGCTGCCCGCCGAGAGAGACCAAGCCTCTGACGTCCATATC 1965  
QY 329 ----- 329  
Db 1966 TCTGCGGTGTGGCAATCTGCTGTGTGCTTTGGGGGTGTCTTTGGGATCTCATC 2025  
QY 329 ----- 329  
Db 2026 AAGGAGGCGCAGCAAGATCCGGAAGTACAGATGCGGAGACTGTGCGAGAAACGGAG 2085  
QY 329 ----- 329

Db 2086 CTGTGGAGCGCGTGTACACCTAGCGGAGCGATGTGCCAAACGAGCGCAGATCGGATCCTG 2145  
QY 329 ----- 329  
Db 2146 AAGAGACGAGCTGAGGAAGGTGAAGGTGCTTGATCTGGCGCTTTTGGCACAGTCTAC 2205  
QY 329 ----- 329  
Db 2206 AAGGCATCTGGATCCCTGATGGGAGATGTGAAAATTCAGTGGCCATCAAGTGTGG 2265  
QY 329 ----- 329  
Db 2266 AGGGAACACATCCCCAAAGCCAAAGAAATCTTAGCAAGCATACGTATGGCT 2325  
QY 329 ----- 329  
Db 2326 GGTGTGGGCTCCCATATGCTCTCCGCTTCTGGGCACTGTGCATCCACGCTGCAG 2385  
QY 329 ----- 329  
Db 2386 CTGTGACACAGCTTATGCCCTATGGCTGCTTATGACCATGTCCGGGAAACCGGGA 2445  
QY 329 ----- 329  
Db 2446 CGCTGGGCTCCAGACCTGCTGACTGTGTATGAGATTGCCAAGGGGATGAGCTAC 2505  
QY 329 ----- 329  
Db 2506 CTGAGGATGTGCGGCTGTACACAGGAGCTTGCCGCTCGGAACGTGTGTCAAGAGT 2565  
QY 329 ----- 329  
Db 2566 CCCAACCATGTCAAAATTACAGACTTGGGGCTGGCTGGCTGTGAGATTGACAGAGACA 2625  
QY 329 ----- 329  
Db 2626 GAGTACCATGAGATGGGGGCAAGTGCCCATCAAGTGTATGGCTGAGTCCATTCTC 2685  
QY 329 ----- 329  
Db 2686 CGCCGGGGTTCAACCCACAGAGTGTGTGAGTTATGTGTGACTGTGTGGAGCTG 2745  
QY 329 ----- 329  
Db 2746 ATGACTTTGGGGCAAACTTACAGATGGATGCCAGCCGGAGATCCCTGACCTGCTG 2805  
QY 329 ----- 329  
Db 2806 GAAAAGGGGAGCGGCTGCCAGCCCCCATCTGCACATTTGATGTACATGATCATG 2865  
QY 329 ----- 329  
Db 2866 GTCAATGTGGATGATGACTGTGAATGTGCGCAAGATTCGGGAGTGTGTCTGA 2925  
QY 329 ----- 329  
Db 2926 TTTCCCGCATGGCCAGGAGCCCCAGGCTTTGTGTGTCATCCAGAAATGAGGACTTGGGC 2985  
QY 329 ----- 329  
Db 2986 CCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGTGAGAGAGATGAGATGGGG 3045  
QY 329 ----- 329  
Db 3046 GACCTGTGATGCTGAGAGATATCTGTACCCAGAGGGCTTCTGTGCCAGACCCT 3105  
QY 330 ----- 347  
Db 3106 GCCCGGGGCTGGGGGCGATGTCCACACAGGACCGCACTCATGTACAGAGGTGGC 3165  
QY 348 GlyIysAspLeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeu 367  
|||||





Db 1066 GTTACAGTCCCAATATCCAGAGTTTGTGCTGCAGAGAGATCTTTGGAGCCTGGCA 1125  
QY 329 ----- 329  
Db 1126 TTTTTCGGGAGAGCTTTGATGGGAGCCAGCCTCCACACTGCCCCGCTCCAGCCAGAG 1185  
QY 329 ----- 329  
Db 1186 CAGCTCAAGTGTGAGACTGTGAGAGATCAGAGTTACTATACATCTCAGCATGG 1245  
QY 329 ----- 329  
Db 1246 CCGAGAGCTGCTGACCTCAGCGTCTTCAGAACTGCAAGTAATCCGGGAGCAATT 1305  
QY 329 ----- 329  
Db 1306 CTGCACATGGGCTACTCGCTGACCTTGCAAGGCTGGGATCAGCTGGGGCTG 1365  
QY 329 ----- 329  
Db 1366 CGCTCACTGAGGGAACGTGGGAGTGGACTGGCCCTCATCCACCATAAACACACCTCTGC 1425  
QY 329 ----- 329  
Db 1426 TTCTGTCACACGCTGCCCTGGGAGCAGCTTTTCGMAACCGCACCAAGCTCTGCTCAC 1485  
QY 329 ----- 329  
Db 1486 ACTGCCAACCGGCCAGAGAGACAGTGTGTGGGCGAGGGCTTGCCCTGCCACCACTGTGC 1545  
QY 329 ----- 329  
Db 1546 GCGCCAGGCACTGCTGGGGGTCCAGGGGCCACCAAGTGTGTCACGTGCAGATTCTT 1605  
QY 329 ----- 329  
Db 1606 CGGGGCCAGAGTGCCTGGAGAGAAATGCCAGTACTGACAGGGCTCCCGAGGAGATGTG 1665  
QY 329 ----- 329  
Db 1666 AATGCCAGGCACTGTTTGGCGTCCACCCTGAGTGTGACCCCAAGATGGCTCAGTGAC 1725  
QY 329 ----- 329  
Db 1726 TGTTTTGACCGAGAGCTGACAGTGTGTGGCTGTGCCACTATTAAGAGCCTCCCTTC 1785  
QY 329 ----- 329  
Db 1786 TGCGTGGCCCCGTGCCCGAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAG 1845  
QY 329 ----- 329  
Db 1846 TTTCCAGATGAGGAGGGCCGATGCCAGCCTTGCCCCCATCACTGACCCACTCTGTGTG 1905  
QY 329 ----- 329  
Db 1906 GACCTGATGACAAGGGTGGCCCCCGAGAGAGAGGAGGCCCTGTGACGTCCATCATC 1965  
QY 329 ----- 329  
Db 1966 TCTGGGCTGTGGCATTTCTGTGCTGTGCTGTGGGGGTGGTCTTTGGGATCCTCATC 2025  
QY 329 ----- 329  
Db 2026 AAGCGACGGCAGCAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGAGGAAACGGAG 2085  
QY 329 ----- 329  
Db 2086 CTGTGTGAGCCGCTGACACCTAGCGGAGGAGATGCCCAACGAGCGAGATGGGATCTTG 2145  
QY 329 ----- 329  
Db 2146 AAGAGACGAGCTGAGGAAGGTGAAGTGTGTGATCTGGCGCTTTTGGCACAGTCTAC 2205

QY 329 ----- 329  
Db 2206 AAGGCATCTGGATCTCCATGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAACTGTG 2265  
QY 329 ----- 329  
Db 2266 AGGAAAAACATTCGCCCAAAAGCAAAAGAAATCTTAGAGAGACATACGTATGGCT 2325  
QY 329 ----- 329  
Db 2326 GGTGTGGCTCCCATATGTCTCCGCTTCTGGGCATCTGCCTGACATCCAGGGTGCAG 2385  
QY 329 ----- 329  
Db 2386 CTGTGTACACAGCTTATGCCCTATGGCTTCCCTTATAGACATGTCCGGGAAACCGCGGA 2445  
QY 329 ----- 329  
Db 2446 CGCTGGCTCCAGAGACCTGCTGAATGTGTATGATGATCCAGGGGATGAGCTAC 2505  
QY 329 ----- 329  
Db 2506 CTGAGAGATGTGGGCTGTACACAGGACTTGGCCCTCGGAAGCTGCTCAAGAGT 2565  
QY 329 ----- 329  
Db 2566 CCCAACATGTCAAAATTAACAGACTTCGGGCTGGCTCGGCTCTGGACATTGACGAGACA 2625  
QY 329 ----- 329  
Db 2626 GAGTACATCAGATGGGGCAAGGTGCCATCAAGTGGAGGCTGGAGTCAATTCTC 2685  
QY 329 ----- 329  
Db 2686 CGCCGGGCTTACCCACAGAGTATGTGTGAGTTATGTGTACTGTGTGGAGCTG 2745  
QY 329 ----- 329  
Db 2746 ATGACTTTTGGGGCCAAACCTTACGATGGGATCCAGCCCGGAGATCCCTACCTGCTG 2805  
QY 329 ----- 329  
Db 2806 GAAAAGGGAGCGGCTGCCAGCCCCCATCTCACCATTGATGTACATGATCATG 2865  
QY 329 ----- 329  
Db 2866 GTCAAAATTTGATGATGACTCTGAATGTGGCCAAAGATTCCGGAGTTGGTGTCTGAA 2925  
QY 329 ----- 329  
Db 2926 TTCTCCCGATGGCCAGGAGCCCGCTTGTGTGTCATCCAGATGAGGACTTGGGC 2985  
QY 329 ----- 329  
Db 2986 CCAGCAGTCCCTTGACAGACACTTTCACGCTCAGCTGAGAGAGAGTACATGGG 3045  
QY 329 ----- 329  
Db 3046 GACCTGTGTGATGCTGAGAGATCTGTACCCAGCAGGGCTTCTGTGTCAGACCT 3105  
QY 330 ----- 347  
Db 3106 GCGCCGGGCTGGGGGCTGTGCTCCAGCGACCGGAGCTCATCTACAGAGTGGC 3165  
QY 348 GtlyAspLeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeu 367  
Db 3166 GGTGGGAGCTTACACTAGAGGCTGGAGCCCTTGAAGAGAGAGGCCCGAGTCTCCAGT 3225  
QY 368 AlaProSerGluGlyAlaGlySerAspAlaPheAspGlyAspLeuGlyMetGlyAla 387  
Db 3226 GCACCTCCGAAGGGGCTGGCTCCGATGTATTGTGTGAGACTGTGGAAATGGGGGCGACGC 3285

```

QY 388 LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAsp 407
DB 3286 AAGGGCTGCMAAGCCCTCCACACATGACCCCGCTTACAGCGGTACAGTGAAGGAC 3345
QY 408 ProThrAlaProLeuProSerGlnThrAspGlyTyrValAlaProLeuThrCysSerPro 427
DB 3346 CCCACAGTACCCCTGCTCTGAGACGTAGTGGCTACGTGGCCCTGACCTGACGAGCC 3405
QY 428 GlnProGlnTyrValAlaGlnProAspValAlaArgProGlnProProSerProArgGly 447
DB 3406 CAGCTTAATATGTGMAACAGCCAGATGTGGCCCGCCCGCTTGGCCCGGAGAGGCG 3465
QY 448 ProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlnArgAlaLysThrLeuSerPro 467
DB 3466 CCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3525
QY 468 GlyLysAsnGlyValValLysAspValAlaPheAlaPheGlyGlyValGluAspProGlu 487
DB 3526 GGGAGAAATGGGGTCTGCAAGAGCTTTTGGCTTGGGGTGGCTGGAGAAACCCGAG 3585
QY 488 TyrLeuThrProGlnGlyAlaAlaAlaProGlnProHisProProAlaPheSerPro 507
DB 3586 TACTTGAACACCCAGGAGAGGCTGCCCTCAGCCCGCCCTCTCTGCTTCAAGCCCA 3645
QY 508 AlaPheAspAsnLeuTyrTyrTPAspGlnAspProProGluArgGlyAlaProProSer 527
DB 3646 GCCTTGCACCACTCTTACTGTGGACAGACCCAGCGGGGGCTCCACCCAGC 3705
QY 528 ThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeuGlnLysLeuAspValPro 546
DB 3706 ACCTTCAAGGACACCTACGGCGAGAGAACCCAGAGTACTGGGTCTGAGAGTCCCA 3762

RESULT 14
US-09-821-883-10 : Sequence 10, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Relner
: APPLICANT: Vidovlc, Damir
: APPLICANT: Graddis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821, 883
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10
: LENGTH: 1437
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER300*rgm-CSF construct
US-09-821-883-10

Alignment Scores:
Pred. No.: 6, 89e-181 Length: 1437
Score: 2188.50 Matches: 426
Percent Similarity: 64.93% Conservative: 22
Best Local Similarity: 61.74% Mismatches: 31
Query Match: 58.53% Indels: 211
DB: 10 Gaps: 4

US-09-821-883-2 (1-690) x US-09-821-883-10 (1-1437)
QY 1 MetLysGlnAlaAlaProLeuLeuAlaAlaAlaAlaSerLeuSerLeuGlnPheLeuPhe 20
DB 1 ATGAGAGCTGACCCCTCTCTGCGCAGGAGCAAGCCTTAGCCCTTGGCTTGTGTT 60
QY 21 LeuLeuPhePheThrPheLeuAspArgSerValLeuAlaLysGlnLeuAlaArgGlyAlaAla 40

```

```

DB 61 CTGCTTTTCTGCTAGACCGAAGTGTACTACCAAGAGTTGGCCGCGGCGCCG 120
QY 41 SerThrGlnAlaCysThrGlnTyrThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
DB 121 TCGACCAAGTGTGACACCGGACAGACATGATGAGTGGGCTCCCTGCACTCCGAGACC 180
QY 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValAlaGlnLysLeuGln 80
DB 181 CACCTGGACATGTCTCCGACCTCTACAGAGGCTGACAGGTGGGAGGAGAAACCTGAA 240
QY 81 LeuThrTyrLeuProThrHisAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100
DB 241 CTCACCTACCTGCTCCACCAATGCCAGCTCTCTCTGAGATATTCAGAGAGTCCAG 300
QY 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
DB 301 GCGTACGTGCTCATCGCTCACACCAAGTGAAGAGAGTCCCATGCGAGAGGCTCGGAT 360
QY 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
DB 361 GTGGAGGACCCAGCTCTTGAAGACACTATGCTCCGCGCTGCTAGACAAATGAGAG 420
QY 141 ProLeuAsnAsnThrProThrProValThrGlyAlaSerProGlnGlyLeuArgGluLeuGln 160
DB 421 CCGCTGAACAATACACCCCTGTCACAGGGGCTCCCGAGAGGCTGCGGAGCTGACAG 480
QY 161 LeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
DB 481 CTTCGAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTGTATTCAGCGGAAACCCACACT 540
QY 181 CysTyrGlnAspThrIleLeuTyrPheAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
DB 541 TGCATCAGCAGACAGATTTGTGMAAGACATCTTCACAAAGAACACAGCTGGCTCTC 600
QY 201 ThrLeuIleAspThrAsnAspSerArgAlaCysHisProCysSerProMetCysLysGly 220
DB 601 ACACGTATACACCAACCCCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAGGC 660
QY 221 SerArgCysTyrPheGlnSerGlnSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240
DB 661 TCCGCGCTGCTGGGAGAGATCTGAGAGATTCAGAGCTGAGCGGCGCACGTGCTGGCC 720
QY 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 260
DB 721 GGTGGCTGTGCTCCGCTCAAGGGGCTCCCTCCACGTACGTGCGCATGTGAGGTGCT 780
QY 261 AlaGlyCysThrGlnProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
DB 781 GCGGCTGACAGGGGCGCCCAAGACACTGTGACTGCTGCTGCTGCTCCACTCAACACAGT 840
QY 281 GlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlnSer 300
DB 841 GGCATGTGAGCTGACCTCCAGCCCTGTCACCTACCAACAGACAGCTTGTAGTCC 900
QY 301 MetProAsnProGlnGlnArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
DB 901 ATGCCCAATCCGAGGGCGCGGTATACATTCGGGCCACGTGTGTGATGCTGCTGCTTAC 960
QY 321 AsnTyrLeuSerThrAspValAlaGlySerGlyAlaGlyLysMetValHisHisArgHisArg 340
DB 961 AACTACTTCTTCTACGACAGCGGATCCGCTAGC----- 993
QY 341 SerSerSerThrArgSerGlnGlyLysAspLeuThrLeuGlnLeuGlnProSerGlnGln 360
DB 993 ----- 993
QY 361 GluAlaProArgSerProLeuAlaProSerGlnGlyAlaGlySerAspValPheAspGly 380
DB 993 ----- 993
QY 381 AspLeuGlnMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
DB 993 ----- 993

```

us-09-821-883-2.p2n.rnpb

Oy	401	LeuGlnAgtGyrSerGluAspProThrValProLeuProSerGluThrAspGlyThrVal	420
Db	993	-----	993
Oy	421	AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln	440
Db	994	-----	1002
Oy	441	ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu	460
Db	1002	-----	1002
Oy	461	ArgAlaLeuThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly	480
Db	1002	-----	1002
Oy	481	GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyValAlaAlaProGlnProHis	500
Db	1002	-----	1002
Oy	501	ProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro	520
Db	1003	-----	1014
Oy	521	GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr	540
Db	1014	-----	1014
Oy	541	LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSerThrGlnPro	560
Db	1015	-----	1053
Oy	561	TrpGlnHisValAsnAlaIleGlnGluAlaArgArgLeuLeuAsnLeuSerArgAspThr	580
Db	1054	TGGAATGCGATGATAGATGCCATCAAGAAGAGCTGAGGCTCTTAATAGACATGGCTGCTG	1113
Oy	581	AlaAlaGluMetAsnGluThrValGluValIleSerGluMetPheAspLeuGlnIlePro	600
Db	1114	GAGAACCAAAAGAAAGCAAGACGTAGACATCATCTCTTAAGATCTTCATCCAGAGCGCG	1173
Oy	601	ThrCysLeuGlnThrArgLeuGluGlyLeuTyrLysGlnGlyLeuArgGlySerLeuThrLys	620
Db	1174	ACATGTGTGCAGACGCCGCTGAAGACTATACAGACAGAGGCTTCAGGGCACTCACCAAA	1233
Oy	621	LeuLysGlyProLeuThrMetMetAlaSerHisTyrLysGlnHisCysArgProThrPro	640
Db	1234	CTCAATGGCGGCTTGACCATGATAGCAGACACTACACAGCAACAGCCCTCCAAACCCG	1293
Oy	641	GluThrSerGysAlaThrGlnIleIleThrPheGluSerPheLysGluAsnLeuLysAsp	660
Db	1294	GAATCACTGTGAAATAGAAAGTACCACCTTGAGAGATTTATAAAGAACTTAAAGGC	1353
Oy	661	PheLeuLeuValIleProPheAspCysTrpGluProValAlaGlnGlyAlaArgProPro	680
Db	1354	TTTCTGTTTGATATCCCTTTTGACTGCTGGAGACGGCGCTCAGAAAGCGCGCACCCCG	1413
Oy	681	ProAlaAlaAlaHisHisHisHisHis 690	
Db	1414	CCG-----	1437

```

: RESUBMIT 15
: US-09-854-356-10
: Sequence 10, Application US/09854356
: Patent No. US200217567A1
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A
: TITLE OF INVENTION: HPR-2/new Fusion Proteins
: FILE REFERENCE: 014056-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356

```

```

CURRENT FILING DATE: 2001-05-09
PRIORITY APPLICATION NUMBER: US 09/493,480
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: US 60/117,976
PRIORITY FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 3955
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: rat HER-2/neu CDNA
NAME/KEY: CDS
LOCATION: (26)..(3799)
OTHER INFORMATION: rat HER-2/neu protein
NAME/KEY: misc_feature
LOCATION: (26)..(1990)
OTHER INFORMATION: extracellular domain (ECD) of rat HER-2/neu
NAME/KEY: misc_feature
LOCATION: (2057)..(3796)
OTHER INFORMATION: intracellular domain (ICD) of rat HER-2/neu
NAME/KEY: misc_feature
LOCATION: (2188)..(3022)
OTHER INFORMATION: kinase domain (KD) of rat HER-2/neu
NAME/KEY: misc_feature
LOCATION: (2999)..(3796)
OTHER INFORMATION: phosphotyrosine domain (PD) of rat HER-2/neu
NAME/KEY: misc_feature
LOCATION: (2999)..(3173)
OTHER INFORMATION: preferred portion of the phosphotyrosine domain
OTHER INFORMATION: (delta PD) of rat HER-2/neu
US-09-854 -356-10

```

Alignment Scores:	
Pred. No.:	3,23e-161
Score:	1968.00
Percent Similarity:	35.87%
Best Local Similarity:	34.02%
Query Match:	52.63%
DB:	9
Gaps:	4.
Length:	3955
Matches:	443
Conservative:	24
Mismatches:	83
Indels:	752
Gaps:	4.

[illegible]

```
Db 395 GACAAATCCCGCCCTCCACCCAGGAGAACCCAGAGGGGCGCGGAGCTGCACACTT 454
QY 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAspProGlnLeuCys 181
Db 455 CGAAGTCTCAGAGAGATCCCGAAGGAGAGATTGTATCCGTGGGAACCTCAGCTTGC 514
QY 182 TyGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAlaLeuThr 201
Db 515 TACAGAGACATGGTTTGTGGAAAGACGCTCTCCGCAAGATATACCAATGCGCTGTC 574
QY 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysGlySer 221
Db 575 GATATACACACCAATCGTCCCGGCTGTCACACTTGTGCCCCCGGCTCAAGACAAAT 634
QY 222 ArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgTrpAlaCysAlaGly 241
Db 635 CACGTGTTGGGGTGAGATCCGGAAAGACTGTCAGATCTTGACTGGCACACTGTGACAGT 694
QY 242 GlyCysAlaAspCysLysGlyProLeuProThrAspCysHisGlnGlnCysAlaAla 261
Db 695 GGTGTGTCGGGTCAAGAGCGCGGCTCCCACTGACTGCTGCATGAGACAGTGCGCGCA 754
QY 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281
Db 755 GGGTGCACGGGGCCCAAGACATCTGACTGCTGCGCTGCCCTCCCACTCAATCATAGTGT 814
QY 282 IleCysGlnLeuHisCysProAlaLeuValThrTyAsnThrAspThrPheGluSerMet 301
Db 815 ATCTGTAGCTGCACCTCCAGCCCTGCTCACCACACACAGACACTTGTGATGCATG 874
QY 302 ProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTyAsn 321
Db 875 CACAACTCAGAGGTGCTGCTACACTTGTGTGCCAGCTGCTGACACCTGCCCCCTACAA 934
QY 322 TyIleuSerThrAspValGlySer----- 329
Db 935 TACCTGTCTACGGAGAGTGGATCTGCACCTGTGTGTCGCCGAATTAACCAAGAGGTC 994
QY 329 ----- 329
Db 995 ACAAGCTGAGACGGAACACAGCGTTGTGAGAAATGCAGCAAGCCCTGTGCTCGAGTGTGC 1054
QY 329 ----- 329
Db 1055 TATGCTCTGGGCATGAGACACTTCGAGGGCGAGGGCCATCACACAGTGCATGTCCAG 1114
QY 329 ----- 329
Db 1115 GAGTTTATGCTGCAGAAAGATCTTTGGAGCCCTGGCATTTTGGCCGAGAGCTTTGAT 1174
QY 329 ----- 329
Db 1175 GGGAGACCCCTCTCCGGCATTTGCTCCGCTGAGGCGCTGAGCAGCTCAAGTTCGAAGCC 1234
QY 329 ----- 329
Db 1235 CTGAGAGAGATCACAGGTACTGTACATCTCAGCATGCGCAGACAGTCTCCGAGACTTC 1294
QY 329 ----- 329
Db 1295 AGTGTCTTCAGAAACCTTCGAATCATTTGGGGAGCGATTCACGATGGCGGTACTCA 1354
QY 329 ----- 329
Db 1355 TTGACACTGCAAGGCCTTGGGGATCCACTCGCTGGGGCTGCGCTACATGCGGAGACTGGGC 1414
QY 329 ----- 329
Db 1415 AGTGAATTGGCTCTGATTCACCGCAAGCCCATCTGTGTTGTACACACTGTACCTTGG 1474
QY 329 ----- 329
Db 1475 GACCAGCTTTCGGGAACCCACATCAGGCCCCGTGCTCCACAGTGGGAACCGCGGGAAG 1534
QY 329 ----- 329
Db 1535 GACTTGTGCGTTCGAGCGGCTTGTGTGTACTGATGTGTGCCACGGGCACTGCTGG 1594
QY 329 ----- 329
Db 1595 GGGCAGAGGCCACCCAGTGTGTCMACTGCAGTCATTTCTTGGGGGCGCAGAGTGTGTG 1654
QY 329 ----- 329
Db 1655 GAGAGTGCCGAGTATGGAAGGGGCTCCCCGGAGTATGTAGTCAAGCCGTGCTG 1714
QY 329 ----- 329
Db 1715 CCGTGTACCCCGAGTGTACAGCTCAAAACAGCTCAGAGACGTGCTTTGATGGAGGCT 1774
QY 329 ----- 329
Db 1775 GATCAGTGTGACGCTGCGCCCACTACAGAGACTGCTCTGTGTGCTGCTGCTGCCCC 1834
QY 329 ----- 329
Db 1835 AGTGTGTGAACCGGACCTCTCTTACATGCCATCTGGAAGTACCGGATGAGAGAGGC 1894
QY 329 ----- 329
Db 1895 ATATGCCAGCGGTGCCCCCATCACTGCACCCAGCTCCTGTGTGATCTGATGAACGAGGC 1954
QY 329 ----- 329
Db 1955 TGCCACAGAGAGAGAGAGCCAGCCCGGTGATCATTCATTCATTCGCACTGTAGAGGCGTC 2014
QY 329 ----- 329
Db 2015 CTGCTGTCTGATCTTATGTGTGTGTGTGTTGGAATCTTAATCAACGAGAGACAGAG 2074
QY 329 ----- 329
Db 2075 ATCCGAAATATAGATGCGTAGGCTGCTGCAGAAACTGATGATGAGAGCCGCTGAGC 2134
QY 329 ----- 329
Db 2135 CCCAGCGAGCAATGCCCAACAGGCTCAGATGCGATCTTAAAGAGAGAGCTTAAGG 2194
QY 329 ----- 329
Db 2195 AAGGTGAAGTGTGTTGATCAGAGCTTTTGGCACTGTCTACAAAGGCAATCTGATCCCA 2254
QY 329 ----- 329
Db 2255 GATGGGAGAAATGTGAATAATCCCGTGTATCAAGGTGTTGAGAGAAACACATCTCCT 2314
QY 329 ----- 329
Db 2315 AAAGCAACAAGAAATTTAGATGAAGCGTATGTGCTGTGTGCTGCTCCGTAT 2374
QY 329 ----- 329
Db 2375 GTGTCCCGCTCTCTGGGCATCTGCTGACATCCACAGTACAGCTGGTACACAGCTTATG 2434
QY 329 ----- 329
Db 2435 CCCTAGGCTGCTCTGTGACATGTCCGGAACACCGAGGTGGCTAGGCTCCAGAGAC 2494
QY 329 ----- 329
Db 2495 CTGCTCAACTGTGTGTGATTCAGATTGCCAAGGGATGAGTACCTGGAGAGCGTGGGCTT 2554
QY 329 ----- 329
Db 2555 GTACACAGGAGACTGTGCTGCCGGAATGTGTAGTCAAGAGTCCCAACACAGTCAAGATT 2614
```

329 ----- 329  
Db 2615 ACAGATTTCGGGCTGGCTGGCTGGACATTGATGAGACAGACTACCATCAGATGGG 2674  
329 ----- 329  
Db 2675 GGCAAGTGGCCATCAATGATGGGATTTGAAATCTATTCTCAGAGCGGGTTCAACCAT 2734  
329 ----- 329  
Db 2735 CAGAGTATGTGTGGAGCTATGAGTACTGTGTGGAGCTGATGACTTTTGGGCCAAA 2794  
329 ----- 329  
Db 2795 CCTTACGATGGAATCCAGCCCGGAGATCCCTGATTGCTGGAGAGGAGAACGCCCTA 2854  
329 ----- 329  
Db 2855 CCTCAGCCTCAATCTGCACCATTTGATGTCTACATGATTATGTCAAATGTTGATGATT 2914  
329 ----- 329  
Db 2915 GACTGTGAATGTGCCCGAGATTCCGGAGTTGGTGCAGAAATTTTACGATATGGCGAGG 2974  
329 ----- 329  
Db 2975 GACCCCCACGCTTTTGTGTGTCATCCAGAACGAGACTTGGCCCATCCAGCCCATGAGC 3034  
329 ----- 329  
Db 3035 AGTACCTTCTACCGTTCTACTGCTGGAAGATGATGACATGGTGACCTGTAGACGCTGAA 3094  
330 ----- -GlyAlaGly 333  
3095 GAGTATCTGGTGGCCCGAGAGGATTTCTCCCGGACCTACCCGACGACCTGGGAGC 3154  
334 MetValHisHisArgHisArgSerSerThrArgSerGlyGlyLeuPheLeu 353  
3155 ACAGCCCATAGAAAGGACCGCAGCTGTCTCCACAGAGTGAGAGTGCTGACCTGACACTG 3214  
354 GlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGlyAla 373  
3215 GCCCTGAGACCTCGGAAAGAGGCCCCCAATCTCCACTGGCTCCCGAAGGGGCT 3274  
374 GlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeu 393  
3275 GGCTCCGATGTGTGATGTGACCTGCAATGGGGGTAAACCAAGGGCTGCAGAGCTTC 3334  
394 ProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeuPro 413  
3335 TCTCCACATGACCTCAGCCCTCTACAGCGGTACAGGAGGAGACCCACATTACTCTGCC 3394  
414 SerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAsn 433  
3395 CCCGAGACTGATGGCTATGTCTCCCTGGCTGAGCCCGACGCCGAGTATGGAAC 3454  
434 GlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAlaArg 453  
3455 CAATCAGAGGTTTCAGCTCAGCTCTTAAACCCAGAGGGTCTCTGCTCCTGTCGG 3514  
454 ProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGlyValVal 473  
3515 CCTGCTGGTGTACTACTAAGAACCCAGACTCTCTCTCGGGAAGATGGGTTGTC 3574  
474 LysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGlnGly 493  
3575 AAAGAGTTTTTGGCTTCGGGGTGTGTGGAAGCCCTGAATACTTACCGAAGAGAA 3634  
494 GlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 513  
3635 GGCACTGCTTCGCCGCCACCTTCTCCTGCTTACGCCACGCTTTGACAAACCTTAT 3694

514 TyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThrPro 533  
3695 TACTGGAGACCAACTCATTCGAGCAGGGGCCCTCCACCAAGTACTTTGAAAGGACCC 3754  
534 ThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSer 553  
3755 ACTGCAGAGAACCTGAGTACTAGCTGATGATGATGATGATGATGATGATGATGATGAT 562  
554 Pro-----SerProSerThrGlnProTrpGlu 562  
3815 CCTGTCTTTCAGAGTGGGGAAGCGCTGACTTGTCTCCATCCGACAAAGCAGGAG 3874  
563 HisVal 564  
3875 AGGCTC 3880

Search completed: May 8, 2003, 14:57:36  
Job time : 285 secs

**THIS PAGE BLANK (USPTO)**

Fri May 9 14:21:07 2003

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 8, 2003, 12:34:18 ; Search time 79 Seconds  
(without alignments)  
2678.569 Million cell updates/secTitle: US-09-821-883-2  
Perfect score: 3739  
Sequence: 1 MRAAPLLARASISLGLF.....EPVCGAPPPAAHHHHH 690Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0Searched: 441362 segs, 153338381 residues  
Total number of hits satisfying chosen parameters: 882724Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summariesCommand line parameters:  
-MODEL=frame+2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO/spool/US09821883/runat\_28042003.104750.14590/app.query.fasta.1.839  
-DB=issued.patents.NA -QFMT=fastap -SUFFIX=p2n.rnt -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdt  
-LIST=45 -DOCAUGEN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09821883.ecgn.1.1.61.etrnat.28042003.104750.14590 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEDQUERY -NEG\_SCORES=0 -MATT=LONGLOG -DEV\_TIMCOUT=120  
-WARN\_TIMCOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database : Issued Patents.NA.\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCITUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2423	64.8	4530	1	US-08-229-515A-9
2	2423	64.8	4530	1	US-08-645-865-9
3	2423	64.8	4530	1	US-09-167-322-4
4	2418	64.7	4473	2	US-09-048-804-1
5	2418	64.7	4473	4	US-09-056-105-26
6	2405.5	64.3	3768	2	US-08-625-101-1
7	2405.5	64.3	3768	2	US-08-356-786-1
8	2258.5	60.4	2385	2	US-09-146-283-3
9	2258.5	60.4	2385	2	US-08-579-823A-3
10	2258.5	60.4	2385	4	US-09-344-195-3
11	1968	52.6	3955	1	US-08-229-515A-14
12	1968	52.6	3955	1	US-08-645-865-14

13	1647	44.0	1872	3	US-08-422-108-2
14	1647	44.0	1872	4	US-08-422-734-2
15	770	20.6	1588	2	US-09-146-283-1
16	770	20.6	1588	3	US-08-579-823A-1
17	770	20.6	1588	4	US-09-344-195-1
18	725.5	19.4	1868	4	US-08-658-863B-1
19	725.5	19.4	1868	4	US-09-676-610B-26
20	724.5	19.4	1593	4	US-09-676-610B-25
21	724.5	19.4	5532	2	US-08-475-035-3
22	724.5	19.4	5532	4	US-09-676-610B-17
23	722.5	19.3	896	6	5200327-2
24	722.5	19.3	900	1	US-08-318-193-7
25	721	19.3	5484	4	US-09-633-580A-3
26	721	19.3	5501	1	US-08-484-438-1
27	721	19.3	5555	1	US-08-484-438-3
28	719	19.2	905	6	5200327-3
29	719	19.2	909	1	US-08-318-193-9
30	714.5	19.1	4905	1	US-07-978-895-3
31	714.5	19.1	4905	1	US-08-473-119-3
32	714.5	19.1	4905	2	US-08-475-352-3
33	714.5	19.1	4975	4	US-09-630-706-3
34	710	19.0	1958	4	US-09-570-454-1
35	693	18.5	822	3	US-08-463-318-69
36	693	18.5	822	3	US-08-463-609A-69
37	693	18.5	822	4	US-08-446-872A-69
38	693	18.5	822	4	US-08-762-227A-69
39	693	18.5	822	5	PCT-US95-01185-69
40	685.5	18.3	903	3	US-08-468-318-66
41	685.5	18.3	903	3	US-08-468-609A-66
42	685.5	18.3	903	4	US-08-446-872A-66
43	685.5	18.3	903	4	US-08-762-227A-66
44	685.5	18.3	903	5	PCT-US95-01185-66
45	677	18.1	392	1	US-08-318-193-1

## ALIGNMENTS

RESULT 1  
US-08-229-515A-9  
Sequence 9, Application US/08229515A  
Patent No. 5518865  
GENERAL INFORMATION:  
APPLICANT: RAZIUDIN  
TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
TITLE OF INVENTION: NEOPLASTIC DISEASE  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESS: NEEDLE & ROSENBERG PC  
STREET: 127 Peachtree Street, Suite 1200  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/229,515A  
FILING DATE: 19 APR 1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: PERRMAN, DAVID G  
REGISTRATION NUMBER: 33,438  
REFERENCE/DOCKET NUMBER: 1414,608  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-688-0770  
TELEFAX: 404-688-9880  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4530 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
us-08-229-515A-9

## Alignment Scores:

Pred. NO.:	6.58e-161	Length:	4530
Score:	2423.00	Matches:	524
Percent Similarity:	41.14%	Conservative:	5
Best Local Similarity:	40.75%	Mismatches:	21
Query Match:	64.80%	Indels:	736
DB:	1	Gaps:	4

US-09-821-883-2 (1-690) x US-08-229-515A-9 (1-4530)

```

OY 2  ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 94  CGCGCCCTCCGACGCGGGTCCAGCGCGGACCATGGGCGCGGAGCCGCGAGTGCACCATG 153
OY 18  PheLeuPheLeuPhePhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
Db 154  GAGCTGGCGGCTTGTGCGCTGG-----GGGCTCTCTCTGCGCCCTTGGCCCCC 204
OY 38  GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 205  GGAGCCCGGACACACCAAGTGTGCACGCGCACAGACATGAAGCTGGGCTCCCTGCGCAGT 264
OY 58  ProGluThrHisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGly 77
Db 265  CCCGAGACCCACCTGAGACATGCTCGCCACCTTACACGAGGCTGCGACAGTGTGTCAGGGA 324
OY 78  AsnLeuGluLeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325  AACCTGGAACTCACCTTACCTGCCACCAATGCCAGCTGTCTCTCTCCAGAGATATCCAG 384
OY 98  GluValGlnGlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
Db 385  GAGGTGCAGGGCTACGAGCTATGCTATGCTACAAACAGTAGAGGAGGTCCACTGCAGAGG 444
OY 118  LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAsp 137
Db 445  CTGGGAGATGTGCGAGGACACCACTGTTTGAGGACAACTATGCCCTGGCGTGTGTAC 504
OY 138  AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 505  AATGAGAGCCCGGTGAACATACCAACCCCTGTACAGGGGCTCCCGCAGAGGCTGCGG 564
OY 158  GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValIleGlnArgAsn 177
Db 565  GAGCTGCAGCTTCCAGAGCTCACAGAGATCTTGAAAGAGGGGCTTGATCCAGCGGAC 624
OY 178  ProGlnLeuCysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln 197
Db 625  CCCAGAGCTGCTACACAGGACAGATTTTGTGAAGGACATCTTCCACAAACACACAG 684
OY 198  LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
Db 685  CTGGCTCTCACACTGATGACACCAACCGCTCTGGGGCTCCACCCCTTCTCCGATG 744
OY 218  CysLysGlySerArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThr 237
Db 745  TGTAAAGGCTCCCGCTGTGGGAGAGAGTCTAGGATTTCTCAAGCTTACCGGCT 804
OY 238  ValCysAlaGlyGlyCysAlaAlaArgCysLysGlyProLeuProThrAspCysHisGlu 257
Db 805  GTCTGTGGCGGTGGCTGTGGCCCTGCAAGGGGACACTGCCCATGTGCTGCATAG 864
OY 258  GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 865  CAGGTGTCTGCGGCTGACACGCGCCCAAGACATGTGACTCTGGGCTGCTCCACTTC 924
OY 278  AsnHisSerGlyIleCysGluLeuHisCysArgProAlaLeuValThrTrpAsnThrAspThr 297

```

```

Db 925  AACCACTGGCAATCTGTGACTGTGACTGCCAGCCCTGGTCCACTTACACACAGACAG 984
OY 298  PheGluSerMetProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAla 317
Db 985  TTTGAGTCCATGCCCAATCCGAGGGCGGTATACATTCGGCGGACGTGTGTACTCTCC 1044
OY 318  CysProTrpAsnTrpLeuSerThrAspValIleSer----- 329
Db 1045  TGTCCCTCAACTACCTTCTTACGAGAGCTGGAGATCTGCACCCCTGTCGCCCTGCAC 1104
OY 329  ----- 329
Db 1105  AACCAAGAGTGACAGACAGAGAGTGAACACAGCGGTGTGAGAAAGTCAGCAGACCCCTGT 1164
OY 329  ----- 329
Db 1165  GCCGAGTGTGCTATATGCTGTGGCATGAGAGACATTGCGAGAGTGAGGGCAGTTACAGT 1224
OY 329  ----- 329
Db 1225  GCCAATATCCAGAGATTGTGTGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCG 1284
OY 329  ----- 329
Db 1285  GAGAGCTTTGATGGGAGCCAGACCTCCACACTGCCCCGCTTCAGGCCAGACAGCTCCAA 1344
OY 329  ----- 329
Db 1345  GTGTTGAGACTGTGAAGAGATCAGAGTTACGTATACATCTGAGCATGSCCGGACAGC 1404
OY 329  ----- 329
Db 1405  CTGCTGACCTACAGGTCCTTCCAGAACCTGCAGATATCCGGGAGCAGAAATCTGCACAA 1464
OY 329  ----- 329
Db 1465  GCGGCTTACTGCTGACCTCTCAAGGGCTGGGCATCACTGCTGGGCTGCGCTCACTG 1524
OY 329  ----- 329
Db 1525  AGGGAATGGGACAGTGCAGCTGGCCCTCATCCACATACACCACCACTGTGCTGTGCAC 1584
OY 329  ----- 329
Db 1585  ACGGTGCCCTGGACACAGCTTTTCGGAACCGGACCAAGCTGTCTCCACACTGCCAAC 1644
OY 329  ----- 329
Db 1645  CGGCCAGAGACGAGTGTGTGGCGAGGGCTGGCTTGCACACAGCTGTGCGCCGAGGG 1704
OY 329  ----- 329
Db 1705  CACTGTGGGTGCCAGGGCCACCCACAGTGTGTCAACTGCACACAGTTCCTTCGGGGCCAG 1764
OY 329  ----- 329
Db 1765  GAGTGCGTGAAGAGATGCCAGTACTGCAAGGGGCTCCCGAGGAGTATGTGAAGCCAGG 1824
OY 329  ----- 329
Db 1825  CACTGTTTGCCCTGCCACCTGAGTGTACGCCAGAAATGGCTAGTACCTGTTTGGGA 1884
OY 329  ----- 329
Db 1885  CCGGAGGTGACAGTGTGTGCTGTGCGCCACTATGAAGACCTCCCTTGTGCGTGCC 1944
OY 329  ----- 329
Db 1945  CGTTCGCCACAGCGGTGTGAACCTGACCTTCTACATGCCCATGTGAAGTTTCCAGAT 2004
OY 329  ----- 329
Db 2005  GAGAGGGGCGCATGCCAGCTTGGCCCATGCACTGCACCCACTCTGTGTGGAACCTGGAT 2064

```





```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/645,865
FILING DATE: 14 May 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PERRYMAN, DAVID G
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414,608
TELEPHONE: 404-688-9880
TELEFAX: 404-688-0770
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4530 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-645-865-9

Alignment Scores:
Pred. No.: 6,58e-161 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 21
Query Match: 64,80% Indels: 736
Gaps: 4

US-09-821-883-2 (1-690) x US-08-645-865-9 (1-4530)
QY 2 ArgAlaIaIaProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 94 CGGCCCTCCACGCCGCTCCACGCCGACGATGGGCCGACGCGATGAGCACCATG 153
QY 18 PheLeuPheLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhe 37
Db 154 GACCTGGCGCGCTGTGCTCCGCTGG-----GGCGCTCCCTCCGCGCTGCCCCC 204
QY 38 GAlaAlaIaIaSerThrGlnValGlyThrGlyThrAspMetLysLeuAlaGluProAlaSer 57
Db 205 GGGAGCCCGGACACCCAGAGTGTGACCGGACACATGAAGCTGGCTCCGCTCCG 264
QY 58 ProGluThrIleuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValAlaGlnGly 77
Db 265 CCGGAGACCCACCTGGACATGCTCCGCGACCTGTACACAGGCGTGCAGGTGCAGGA 324
QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325 AACCTGAACTCACTCACTGCGCCACCAATGCCAGCTGTCTCCGTCGAGATATCCAG 384
QY 98 GAlaValGlnGlyTyrValLeuIleAlaHisAsnGlnValAlaGlnValAlaProLeuGlnArg 117
Db 385 GAGGTGAGGGGCTACGTCGCTCATGCTGACACACCAAGTGAAGGAGGCTCCACCTG 444
QY 118 LeuArgIleValAlaArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValAlaAsp 137
Db 445 CTGCGGATTTGGGAGGACCCAGCTCTTTGAGAGACACATATGCCCTGGCGCTGCTAGAC 504
QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 505 AATGGAGACCCGCTGAACATATACACCCCTGTCAAGAGGCGCTCCAGAGAGGCTCGG 564
QY 158 GlnLeuGlnLeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnAsnArg 177
Db 565 GACCTGAGGCTTGAAGCCCTACAGAGATCTTAAAGAGAGGAGGCTTTGATCCAGCGGAC 624
QY 178 ProGlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGln 197
Db 625 CCGCAGCTGTCTACACGAGACAGATTTGTGAGAGGACATCTTCCACAGAAACACAG 684
```

```

QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
Db 685 CTGCTCTCACTGATGATAGACACCAACCGCTCTGGCGCTGCCACCCCTGTCCGATG 744
QY 218 CysLysGlySerArgGlyTyrPheGlyLeuSerSerLysProCysGlnSerLeuThrArgThr 237
Db 745 TGTAAAGGCTCCCGCTCTGGGAGAGATTTGTAGATTTGTGAGACCTGACCGGCACT 804
QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGly 257
Db 805 GTCTGTCCGCTGGCTGTGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 864
QY 258 GlnCysAlaAlaGlyCysThrGlnGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 865 CAGTGTCTGCGGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 924
QY 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
Db 925 AACCAAGTGGCATCTGTGAGCTGCACTGCCACCCCTGTGACCTTCAACACAGACAGC 984
QY 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
Db 985 TTGTAGTCCATGCCCAATCCCGAGGCGCGGTATATCATTCGCGCCAGCTGTGTACTGCC 1044
QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
Db 1045 TGTCCCTCAACAATCACTTTCTACGAGAGCTGGATCTGTACACCTGTGTGCCCCCTGCAC 1104
QY 329 ----- 329
Db 1105 AACCAAGAGGTGACAGCAGAGATGAGAACACAGCGGTGTGAGAAAGTGCACCAACGCTGT 1164
QY 329 ----- 329
Db 1165 GCCGAGTGTGCTATGCTGTGGGATGAGACACTTGGAGAGGTGAGGAGGAGTATACAGT 1224
QY 329 ----- 329
Db 1225 GCCAAATCCAGAGATTGTGCTGCTGCAGAGAAAGATCTTTGGAGCCCTGGCATTTCTG 1284
QY 329 ----- 329
Db 1285 GAGAGCTTTATGAGGAGGACCAAGCTCCAACTGCCCCCTCCAGGACAGAGAGCTCCA 1344
QY 329 ----- 329
Db 1345 GTGTTGAGACTGTGGAAGAGATCACAGGTTACCTATATCATTCAGCATGAGCGCGACAGC 1404
QY 329 ----- 329
Db 1405 CTGCTGACCTCAGCGCTTTCAGAACTCAGAACTAATCCGGGAGAGAAATTTGACAAAT 1464
QY 329 ----- 329
Db 1465 GCGGCTTACTGCTGACCTTGCAGAGGCTGGGATCAGCTGCTGGGGCTGCGCTCACTG 1524
QY 329 ----- 329
Db 1525 AGGGAAGTGGGAGTGAAGTGGCGCTTATCCACCATTAACACCACTCTGTCTGTGCTGAC 1584
QY 329 ----- 329
Db 1585 ACGGTGCTGGGAGACAGCTTTTTCGAAACCCGACCACTGTCTGTCACTGTGCCAAC 1644
QY 329 ----- 329
Db 1645 CGGCCAGAGAGAGATGTGTGGGAGAGGCTGTGCTGCCACAGCTGTGCTGCCAGAGG 1704
QY 329 ----- 329
Db 1705 CACTGCTGTGGGTCCAGGCGCCACCAAGTGTGTCACTGACGAGCAGTTCTTTCGGGGGCGAG 1764
```

QY 329 ..... 329  
Db 1765 GAGTCGTGAGAGATGCCAGTACTGACAGGGGCTCCCGAGGAGTATGTAAATCCAGG 1824  
QY 329 ..... 329  
Db 1825 CACTGTTTGCCTGCCACCCTGAGTGTAGCCCCAGAAATGGCTCAGTACCTGTTTGA 1884  
QY 329 ..... 329  
Db 1885 CCGAGGCTGACCACTGTGTGCTGTGCCCACTAAGAGACCTCCCTTCTGCTGGCC 1944  
QY 329 ..... 329  
Db 1945 CGCTGCCCGAGGGGTGTGAACCTGACCTTCCTACATGCCCATCTGGAAGTTTCCAGAT 2004  
QY 329 ..... 329  
Db 2005 GAGGAGGGCGCATGCCAGCCTTGCCCATCAACTGCACCACCTCCTGTGTGAGACTGGAT 2064  
QY 329 ..... 329  
Db 2065 GACAGAGGCTGCCCGCCGAGAGAGACCAAGCCCTGTAGCTCCATGCTCTCTGGGTG 2124  
QY 329 ..... 329  
Db 2125 GTTGGCATTTCTGCTGTGTGTGTTGGGGGTGTCTTTGGGATCCTCATCAAGCGACGG 2184  
QY 329 ..... 329  
Db 2185 CAGCAGAAAGATCCGGAGTACACGATGGGAGACTGTGCAGAAACGAGCTGGTGAG 2244  
QY 329 ..... 329  
Db 2245 CCGCTGACACCTAGCGGAGCGGATGCCCAACGAGCGCAGATCGGATCCTGAAGAGAG 2304  
QY 329 ..... 329  
Db 2305 GAGCTGAGGAAGTGAAGTCTTGATCTGGCCTTTTGACACATCTACAAGGCGATC 2364  
QY 329 ..... 329  
Db 2365 TGGATCCCTGATGGGAGAAATGTAAATTCAGTGGCCATCAAGTGTGAGGAAAC 2424  
QY 329 ..... 329  
Db 2425 ACATCCCCCAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGCTGTGGGC 2484  
QY 329 ..... 329  
Db 2485 TCCCATATGTCTCCGCTTCTGGGCATCTGCTGACATCCAGGTCAGCTGTGACA 2544  
QY 329 ..... 329  
Db 2545 CAGCTTATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGACGCTGGGC 2604  
QY 329 ..... 329  
Db 2605 TCCCAAGACCTGCTGAACCTGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGAT 2664  
QY 329 ..... 329  
Db 2665 GTGCGGCTGTACACAGGACTTGGCGCGTGGAAAGTGTGTAAGATCCCAACAT 2724  
QY 329 ..... 329  
Db 2725 GTCAAAATTACAGACTTGGGCTGGCTGGCTGTGACATTGACAGAGACAGATACAT 2784  
QY 329 ..... 329  
Db 2785 GCAGATGGGGGCAAGGTCCCATCAAGTGGATGGCGTGGAGTCCATTTCGCGCGCGG 2844  
QY 329 ..... 329

Db 2845 TTACCCACCAAGATGATGTGTGAGATTATGGTGTACTGTGTGGAGCTGATGACTTTT 2904  
QY 329 ..... 329  
Db 2905 GGGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCCTGACTGCTGAAAAAGGGG 2964  
QY 329 ..... 329  
Db 2965 GAGCGGCTGCCCGACCCCGCATGTCACCATATGATGTCTACATGATGATGTCAAATGT 3024  
QY 329 ..... 329  
Db 3025 TGGATGATTACTGTGAATGTGCGCAAGATTCCGGAGTGTGTGTGAATTCTCCCGC 3084  
QY 329 ..... 329  
Db 3085 ATGCCAGGAGACCCCGACCGCTTTGTGTGATTCAGAAATGAGACTTGGGCCACCGCAT 3144  
QY 329 ..... 329  
Db 3145 CCCTTGAGACGACCTTCTACCGCTCAGCTGTGAGAGACATGACATGGGGAGCTGGTG 3204  
QY 330 ..... 330  
Db 3205 GATGCTGAGAGATATCTGTGATCCCGAGGCGCTTCTGTCTCAGACCTGCCCGCGGC 3264  
QY 331 AlaGlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyLys 350  
Db 3265 GCTGGGCGATGTGCACACAGGACCGCACCTCATCTATCCAGAGAGTGGCGTGGGAC 3324  
QY 351 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer 370  
Db 3325 CTGACACTAGGGCTGAGCCCTCTGAAGAGAGAGGCCCGCCAGTGTCCACTGGCACCTCC 3384  
QY 371 GluGlyAlaGlySerAspValPheAspGlyLysLeuGlyMetGlyAlaAlaLysGlyLeu 390  
Db 3385 GAAGGGCGTGGCTCCGATGATTTGATGTGCTGCGGGAATGGGCGACCCAAAGGGCTG 3444  
QY 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410  
Db 3445 CAAAGCTCCCGACACATGACCCCGACCTCTACAGCGGTACGTAGAGAGACCCACAGTA 3504  
QY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 430  
Db 3505 CCCCCTGCCCTGAGACATGATGGTACGTGGCCCCCTGACCTGCAGCCCCAGCTGTGCT 3564  
QY 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 450  
Db 3565 TATGTGAACCAAGCAGATGTTGCGCCCGACCCCTTGGCCCGAGAGGGCCCTTGCT 3624  
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluAlaGlyAlaLysThrLeuSerProGlyLysAsn 470  
Db 3625 GCTGCCCGACCTGTGTGGCCACTGTGAAAGGSCAAAGACTCTCCCGACGAGAGAT 3684  
QY 471 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490  
Db 3685 GGGGTGTCGAAGAGCTTTTCCCTTGGGGGTGCGGTGAGAAACCCGAGTACTTGACA 3744  
QY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510  
Db 3745 CCCCAGGGAGGAGCTGCCCTCAGCCCAACCTCTCCCTGAGCCCTTCCAGCCCTTCGAC 3804  
QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530  
Db 3805 AACCTCTATTACTGTGGACAGACCCACAGGCGGGGGCTCCACCCAGCATCTTCAAA 3864  
QY 531 GlyThrProThrAlaGluAsnProGluTyrTyrLeuGlyLeuAspValProAlaAlaPro 550  
Db 3865 GGGACACCTACGGCAGAGAACCAAGTACCTGGGTCTGGAGCTGCCA-----GTGTA 3918  
QY 551 AlaArgSerProSerPro 556  
||| |||||

Db 3919 ACCAGAGCCAGCTCCG 3936

RESULT 3

US-09-167-322-4

Sequence 4, Application US/09167322

Patent No. 6365151

GENERAL INFORMATION:

APPLICANT: Allegheny University of the Health Sciences, Halpern, Michael S.

England, James M.

TITLE OF INVENTION: CANCER VACCINE

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.

STREET: Suite 1800, Two Penn Center Plaza

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19102

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/167,322

FILING DATE: 07-Oct-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/00582

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.

REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 7933-33 PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 4530 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-167-322-4

Alignment Scores:

Pred. No.:	6,58e-161	Length:	4530
Score:	2423.00	Matches:	524
Percent Similarity:	41.14%	Conservative:	5
Best Local Similarity:	40.75%	Mismatches:	21
Query Match:	64.80%	Indels:	736
DB:	4	Gaps:	4

US-09-821-883-2 (1-690) x US-09-167-322-4 (1-4530)

QY 2 ArgAlaAlaProLeuLeuLeuAlaAlaAla-----AlaSerLeuSerLeuGly 17

Db 94 CGGCCCTCCAGCGCGGTCACGCGAGCATGCGGCGCGAGCGAGGACAGCATG 153

QY 18 PheLeuPheLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhe 37

Db 154 GAGCTGGCGGCTTGTCCCTGG-----GGGCTCTCTCTCTCTCTCTCTCTCTCT 204

QY 38 GAlAlAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57

Db 205 GGAGCCCGGAGACCCCAAGTGTGACCGGACGACATGAAGCTGGCGCTCCGCGAGT 264

QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrgInglyCysGlnValValGlnGly 77

Db 265 CCGGAGACCCACCTCGACATGCTCCGCCACTCTACAGGGCTGCCAGGTGTGTCAGGA 324

QY 78 AsnLeuGluLeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97

Db 325 AACCTGGAACTACACTACCTGCGCCACCAATGGCCAGCTGTCTCTCTCGAGGATATCCAG 384

QY 98 GluValGlnGlyTrpValIleLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117

Db 385 GAGGTGACGGCTACAGTGCATCTGCTACACCAACCAATGAGGCGGTGCCACTGCAAGG 444

QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAsp 137

Db 445 CTGGGATGTGGCAGGCGACCCAGCTCTTGAGGACAACTATGCGCTGGCGCTCTGAC 504

QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157

Db 505 AATGAGACCCCGCTGACAAATATACCACTCTGTACAGGGCGCTCCAGAGAGCTCGCG 564

QY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValIleLeuIleGlnArgAsn 177

Db 565 GAGCTGACGCTTGCACACCTCAGAGATCTTGAAAGAGGGGTCTTGATTCACCGGAGAC 624

QY 178 ProGlnLeuCysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGln 197

Db 625 CCCAGCTCTGCTACAGGACAGCATTTTGTGAAAGACATCTTCCACAAAGAACACAG 684

QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217

Db 685 CTGGCTCTCACACTGATGATACACCAACCGCTCTGGCGCTCCACCCCTGTCTCCATG 744

QY 218 CysLysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThr 237

Db 745 TGTAAAGGCTCCGCTGCTGGGAGAGATGTTGAGATGTGTACAGCTTGACCGGCACT 804

QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu 257

Db 805 GTCTGTCGGGTGCTGCTGCGCGCTGCAAGGGCGCACTGCCACTGACTGCTGCAATGAG 864

QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277

Db 865 CAGTGTCTCCGCTGCTGCT 924

QY 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThr 297

Db 925 AACCAAGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTTCAACACAGACAG 984

QY 298 PheGluSerMetProAsnProGluGlyArgTrpPheGlyAlaSerCysValThrAla 317

Db 985 TTTCAGTCCATGCCCAATCCCGAGGCGCGGATATTCAGGCGCGCACTGTGACTGCC 1044

QY 318 CysProTrpAsnTrpLeuSerThrAspValGlySer----- 329

Db 1045 TGTCCCTTACAACTACTTCTTACGAGGTGGATCTGCACCCCTGCTGCCCTGCAC 1104

QY 329 ----- 329

Db 1105 AACCAAGAGTGCAGCAGAGATGGAACACAGCGGTGTGAGAGTGCACAGCCCTGT 1164

QY 329 ----- 329

Db 1165 GCCGAGTGTGATGTGTGTGGCATGAGCACTTGGAGAGTGAGGCGAGTTACAGT 1224

QY 329 ----- 329

Db 1225 GCCAATATCCAGAGTTGTGTGCTGCTGCAGAGATCTTTGGAGCCTGGCATTTCTGCG 1284

QY 329 ----- 329

Db 1285 GAGAGCTTTGATGGGAGCCAGCGCTCCAAACATGCCCCGCTCCAGGACAGACTCCAA 1344

QY 329 ----- 329

Db 1345 GTGTTTGAAGACTCTGAAGAGATCAGAGTTTACTATATCTCAGCATGAGCGGAGAGC 1404

QY 329 ----- 329

Db 1405 CTGCTGACTCAGGCTCTTCAGAACTGCAAGTAATCCGGGAGCAATTTCTGCACAAAT 1464  
 Oy 329 ----- 329  
 Db 1465 GGGCCCTACTGCTGACCTTCGCAAGGGCTGGGCATCAGCTGGGGCTGGCTGACTG 1524  
 Oy 329 ----- 329  
 Db 1525 AGGGAAGTGGGCAATGAGACTGGCCCTCATTCACATTAACACCACCTGCTGCTGAC 1584  
 Oy 329 ----- 329  
 Db 1585 ACGGTGCCCTGGGACACAGCTCTTTGGAAACCGCACAAAGCTCTGCTCACACTGCCAAC 1644  
 Oy 329 ----- 329  
 Db 1645 CGGCCAGAGACAGTGTGTGGGGAGGGCCCTGGCTGCCACAGCTGTGCCGCCGAGGG 1704  
 Oy 329 ----- 329  
 Db 1705 CACTGCTGGGGTCCAGGGGCCACCAACGATGTGTCAACTGACGCAATTCCTTCGGGGCCAG 1764  
 Oy 329 ----- 329  
 Db 1765 GAGTGCCTGAGAGAAATGCCGAGTACTGACAGGGGCTCCCCAGGAGTATGTGAATGCCAG 1824  
 Oy 329 ----- 329  
 Db 1825 CACTGTTCCTGCTGCCACCTGAGTGTGACGCCAGAAATGGCTAGTACCTGTTTGA 1884  
 Oy 329 ----- 329  
 Db 1885 CCGGAGGCTGACCAAGTGTGTGCTGTGCCACTATAAGAGACCTCCCTTCTGCTGGCC 1944  
 Oy 329 ----- 329  
 Db 1945 CGCTGCCAGCGGTGTGAAACCTGACCTTCCTACATGACCATGTGGAAGTTTCCAGAT 2004  
 Oy 329 ----- 329  
 Db 2005 GAGGAGGGCGATGCCAGCCTTGCCTCCATCACTGACACCACCTCTGTGTGACCTGGAT 2064  
 Oy 329 ----- 329  
 Db 2065 GACAAGGGCTGCCCGCCGAGAGAGACCAAGCCTCTGAGCTCCATGCTCTCTCGGTG 2124  
 Oy 329 ----- 329  
 Db 2125 GTTGGCAATTCGTGCTGTGTGTCTTTGGGGGTGTCTTTGGGATCCTCATCAAGCGACGG 2184  
 Oy 329 ----- 329  
 Db 2185 CAGCAGAAGATCCGGAAGTACAGATGCGGAGACTGTGACGAAACGAGAGCTGTGTGAG 2244  
 Oy 329 ----- 329  
 Db 2245 CCGGTGACACCTAGCGGAGCGATGCCCAACGAGCGCAATCGGATCTCTGAAGAAGAG 2304  
 Oy 329 ----- 329  
 Db 2305 GAGCTGAGGAAGTGAAGTGTGTGATCTGGCGCTTTTGGACAGTCTACAAAGGCAATC 2364  
 Oy 329 ----- 329  
 Db 2365 TGGATCCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAGAGTGTGAGGGAAGAA 2424  
 Oy 329 ----- 329  
 Db 2425 ACATCCCCCAAGCAACAAAGAAATCTTAGACGAAGCATACGTATGGCTGTGTGGG 2484  
 Oy 329 ----- 329

Db 2485 TCCCCATATGTCCTCCGCTCTTGCGGACTGTGCTGACATCCAGGTGCACTGTGTACA 2544  
 Oy 329 ----- 329  
 Db 2545 CAGCTTATGCCCTATGGCTGCTCTTAGACCATGTCCGGAAACCGGACGCTGGGC 2604  
 Oy 329 ----- 329  
 Db 2605 TCCAGGACCTGTGTAAGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGAT 2664  
 Oy 329 ----- 329  
 Db 2665 GTGCGCTGTACACAGGAGACTTGGCCGCTGGAAAGCTGTGTCAAGATCCCAACCAT 2724  
 Oy 329 ----- 329  
 Db 2725 GTCAAAATTACAGACTTGGGCTGGCTGGCTGCTGACATTTGACGACAGATACCAT 2784  
 Oy 329 ----- 329  
 Db 2785 GCAGATGGGGGCAAGTGGCCCATCAAGTGAATGGCGCTGAGTCCATTCTCCGCTGGCG 2844  
 Oy 329 ----- 329  
 Db 2845 TTCACCAACAGAGTATGTGTGAATTATGTGTGACTGTGTGGAGACTGACTTTT 2904  
 Oy 329 ----- 329  
 Db 2905 GGGCCAAACCTTACAGATGGATCCAGCCGAGGATCCCTGACCTGTGAAAGAGG 2964  
 Oy 329 ----- 329  
 Db 2965 GAGCGGCTGCCACAGCCGCCCATCTGCAACATTAATGTCTACATGATGTCAAATGT 3024  
 Oy 329 ----- 329  
 Db 3025 TGGATGATTAATCTGAATGTGCGCAAGATTCCGGAGTGTGTGTAATTCCTCCGC 3084  
 Oy 329 ----- 329  
 Db 3085 ATGCCAGGAGACCCAGCGCTTGTGTGTCATCCAGAAATGAGACTTGGGCCACGCACT 3144  
 Oy 329 ----- 329  
 Db 3145 CCCTTGAGACAGCCTTCTACCGCTCACTGCTGAGAGACGATGACATGGGGAGCTGGT 3204  
 Oy 330 ----- 330  
 Db 3205 GATGCTGAGAGTATCTGTACCCAGCAGGCGCTTCTGTCCAGACCTGCCCGGGC 3264  
 Oy 331 AAGTGLYMeTValHisHisArgHisArgSerSerThrArgSerGlyGlyLys 330  
 Db 3265 GCTGGGGCATGTGTCACACAGCGACGCACTCATCTACAGAGAGTGGCGTGGGAC 3324  
 Oy 331 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer 370  
 Db 3325 CTGACACTAGGGCTGAGACCCCTCTGAAGAGAGGCCCCAGAGTCTCCACTGGCACCTCC 3384  
 Oy 3385 GAAAGGGCTGGCTCCGAGTATTGATGTGACCTGGGAATGGGGGAGCCAAAGGGGCTG 3444  
 Db 3391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410  
 Oy 3445 CAAGCCTCCACACATGACCCAGCCTCTACAGCGGTACAGTGAAGACCCCAAGTA 3504  
 Db 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 430  
 Oy 3505 CCGCTGCCCTGTGACATGATGAGGTCAAGTGGCCCCCTGACCTGACGCCCAAGCTGAA 3564  
 Db 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyGlyProLeuPro 450  
 Oy 3565 TATGTGAACCAAGCATGATGTGGGCCCGAGCCCTTGGCCCGAGAGGGCCCTCTGCT 3624

QY 451 AAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn 470  
 DB 3625 GCTGCCGACCTGCTGTGCACCTCTGAAAGGCCAAGACTCTCTCCACGAGGAACAT 3664  
 QY 471 GAlaValAlaLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490  
 DB 3685 GGGGTCTCTCAAGACGTTTGTGCTTGGGGGTCCCTGAGAACCCCGAGTACTTGACA 3744  
 QY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510  
 DB 3745 CCCCAGGAGAGAGCTGCCCTCAGCCCCCAGCTCTCTCTGCTTCCAGCCAGCTTGAC 3804  
 QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530  
 DB 3805 AACCTCTATTACTGGACGACGACCCACAGAGCGGGGCTCCACCCAGACCTTCAAA 3864  
 QY 531 GAlaThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaPro 550  
 DB 3865 GGGACACCTAGCGGAGAAACCCAGAGTACCTGGTCTGGACGTGCA-----GTGTGA 3918  
 QY 551 AlaArgSerProSerPro 556  
 DB 3919 ACCAGAGGCCAGTCCG 3936  
 RESULT 4  
 US-09-048-804-1  
 ; Sequence 1, Application US/09048804  
 ; Patent No. 5968748  
 ; GENERAL INFORMATION:  
 ; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters  
 ; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748rls LLP  
 ; STREET: One Liberty Place - 46th Floor  
 ; CITY: Philadelphia  
 ; STATE: PA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 19103  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 1.44 Mb diskette  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/048,804  
 ; FILING DATE: Herewith  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Paul K. Legard  
 ; REGISTRATION NUMBER: 38,534  
 ; REFERENCE/DOCKET NUMBER: ISIS-2913  
 ; TELEPHONE: (215) 568-3100  
 ; TELEFAX: (215) 568-3439  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4473 base pairs  
 ; TYPE: Nucleic Acid  
 ; STRANDEDNESS: Single  
 ; TOPOLOGY: Unknown  
 ; ANTI-SENSE: NO  
 ; US-09-048-804-1  
 Alignment Scores:  
 Pred. No.: 1.45e-160 Length: 4473  
 Score: 2418.00 Matches: 523  
 Percent Similarity: 41.06% Conservative: 5  
 Best Local Similarity: 40.67% Mismatches: 22  
 Query Match: 64.67% Gaps: 736  
 DB: 2

US-09-821-883-2 (1-690) x US-09-048-804-1 (1-4473)  
 QY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17  
 DB 118 CGGCCCTCCACGCGGGGTCCAGCCGACCATGGGCGCGGAGCCGAGTGAAGACCATG 177  
 QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37  
 DB 178 GAGCTGGCGGGCTTGTCCCGCTGG-----GGGTCTCTCTCGCCCTTGGCCCCC 228  
 QY 38 GAlaAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57  
 DB 229 GGAACCCCGAGACACCAAGTGTGCACCGGACAGACACTGAAGCTGCGGCTCCCTCCAGT 288  
 QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77  
 DB 289 CCGGAGACCCACCTGGACATGCTCCGACCTCTACACAGGGCTGCGAGTGTCCAGGA 348  
 QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97  
 DB 349 AACCTGGACCTCACTACCTGACCCACCAATGCCAGCTGTCTCTTCCAGAGATATCCAG 408  
 QY 98 GluValGlnGlyTyrValIleLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117  
 DB 409 GAGGTGAGGGGCTAGCTGCTATGCTGCACACCAAGTGAAGGAGGTCCAGCTGAGAGG 468  
 QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137  
 DB 469 CTGGGATTTGTGCGAGGACCCACCTCTTGGAGACAACTATGCTCCGCGCTGCTAGAC 528  
 QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157  
 DB 529 AATGAGACACCCGCTGAGCAATACACCCCTGTCCAGAGGGCTCCACGAGAGGCTCTCGG 588  
 QY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsn 177  
 DB 589 GAGTGCAGCTTGAGAACCTCAGACAGATCTTGAAGAGAGGGCTTGATCCAGCGGAC 648  
 QY 178 ProGlnLeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGln 197  
 DB 649 CCCAGCTGTGCTACAGAGACAGATTTTGTGAAGAGACATCTTCCACAGAACACACAG 708  
 QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217  
 DB 709 CTGGCTCTCACACTGATGACACACCAACCGCTCTGCGGCTGCGCACCCCTGTCTCCAGT 768  
 QY 218 CysLysGlySerArgCysTrpGlyLysSerSerGlnAspCysGlnSerLeuThrArgThr 237  
 DB 769 TGTAAAGGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTCAAGCCTGACCGCAGCT 828  
 QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu 257  
 DB 829 GTGTGTCGGGTGGCTGTGCCCGCTGCAGAGGGGCACTGGCCACTGACTGTGTCAGTACG 888  
 QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277  
 DB 889 CAGTGTCTCCGCGCTGCACGAGGGGCCCCAACACACTGTGCTGCGGCTGCTGCTGCTGCT 948  
 QY 278 AsnHisSerGlyLysGlyGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297  
 DB 949 AACCCAGTGCAGCTGTGAGCTGACCTGACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTG 1008  
 QY 298 PheGluSerMetProAsnProGluGluArgTyrThrPheGlyAlaSerCysValThrAla 317  
 DB 1009 TTGTAGTCCATGCCCAATCCCGAGGGCGGTATATCATTCGCGCCAGCTGTGTGACTGCC 1068  
 QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329  
 DB 1069 TGTCCCTACAACTACCTTTCTACGAGGTGGATCTCGACACCTCGTCTGCCCTGCTGAC 1128  
 QY 329 ----- 329

Db 1129 AACCAAGAGTGACAGACAGAGGATGAAACACAGCGGTGTGAGAAATGACAGCAAGCCCTGT 1188  
Oy 329 ----- 329  
Db 1189 GCCCGAGTGTCTATGTCTGTGGCATGGAGCACTTTCGAGAGGTGAGGGCAGTTACCACT 1248  
Oy 329 ----- 329  
Db 1249 GCCAATATCCAGAGATTGTGCTGTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGTGCG 1308  
Oy 329 ----- 329  
Db 1309 GAGACCTTTGATGGGAGCCAGACCTCCAACACTGCCCGCTCCAGCAGAGACTCCAA 1368  
Oy 329 ----- 329  
Db 1369 GTGTTGAGACTCTGGAGAGATCAGAGTTACCTATACATCTCAGCATGCCCCGACAGC 1428  
Oy 329 ----- 329  
Db 1429 CTGCCTGACCTCAGCGTCTTCCAGAACCTGCAGATTAATCCGGGGAGCAATTTCTGCACAAT 1488  
Oy 329 ----- 329  
Db 1489 GGGCCCTACTGCTGACCCCTGCAGAGGCTGGGCAATCAGCTGGGCTGGGCTCAGCTG 1548  
Oy 329 ----- 329  
Db 1549 AGGGAAGTGGGAGTGGAGCTGCGCCTCATCCACATTAACACCACCTCTGCTTGTGCAC 1608  
Oy 329 ----- 329  
Db 1609 ACGTGCCCTGGGACACACTCTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCAAC 1668  
Oy 329 ----- 329  
Db 1669 CGGCGAGAGACAGTGTGTGGGCGAGGGCTGCGCTGCACACAGCTGTGCGCCGAGG 1728  
Oy 329 ----- 329  
Db 1729 CACTGCTGGGCTCAGAGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTGGGGCCAG 1788  
Oy 329 ----- 329  
Db 1789 GAGTGGAGAGATGCCAGTACTGACAGGGGCTCCGAGGAGTATGTGAATCCAGG 1848  
Oy 329 ----- 329  
Db 1849 CACTGTTGCCCTGCCACCTGAGTGTAGCCCCAGAAATGGCTCAGTCACTGTTTGA 1908  
Oy 329 ----- 329  
Db 1909 CCGGAGCTGACCACTGTGTGTGCTGTGCCACTATTAAGACCTCCCTTCTGCGTGCC 1968  
Oy 329 ----- 329  
Db 1969 CGCTCCCCAGGAGTGTGAACCTGACCTCTCTACTATGCCATGTGAAGTTTCCAGAT 2028  
Oy 329 ----- 329  
Db 2029 GAGGAGGCGCATGCGACGCTTGCCTCACTACCTCCTGTTGTGAGCTGAT 2088  
Oy 329 ----- 329  
Db 2089 GACAAGGCTGCCCGCGAGCAGAGAGCCACTCTGAGCTCATCATCTTCTCGGGTG 2148  
Oy 329 ----- 329  
Db 2149 GTTGCAATTCTGTGTGTGTGTTTGGGGGTGTCTTTGGGATCCTCATCAAGGAGCG 2208  
Oy 329 ----- 329  
Db 2209 CAGCAGAAGATCCGGAAGTACAGATGCGAGACTGTCTCAGAAACGAGCTGTGTGAG 2268

Oy 329 ----- 329  
Db 2269 CCGCTGACACTAGCGAGCGATGCCCAACAGGCGCAGATGCCGATCCTGAAGAGACG 2328  
Oy 329 ----- 329  
Db 2329 GAGCTGAGAGATGAAGTGTCTGATCTGGCGCTTTGGCAGAGTCTACAGGCGATC 2388  
Oy 329 ----- 329  
Db 2389 TGGATCCCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAGTTTGAGGAAAAAC 2448  
Oy 329 ----- 329  
Db 2449 ACATCCCCCAAGCCACAAAGAAATCTTAGACGAGCATACGTATGCTGTGTGCGC 2508  
Oy 329 ----- 329  
Db 2509 TCCCATATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCAGGTCACACTGTGACA 2568  
Oy 329 ----- 329  
Db 2569 CAGCTTATGCCCTATGCTGCTGCTCTTAGACCATGTCCGGAAAAACCGCGAGCCTGGGC 2628  
Oy 329 ----- 329  
Db 2629 TCCAGACCTGCTGAAGTGTGTATGTCAGATTGCCAAGGGAGTACCTGAGAGAT 2688  
Oy 329 ----- 329  
Db 2689 GTGGGCTGTACACAGGAGTGTGGCGCTCGAAGCTGTGTCAAGTCCCAACCAT 2748  
Oy 329 ----- 329  
Db 2749 GTCAAAATTACAGACTTGGGCTGCTGCTGTGACATTGACGAGACAGATACCAT 2808  
Oy 329 ----- 329  
Db 2809 GCAGATGGGGGCAAGTCCCATCAAGTGATGGGCTGAGTCCATTCTCGCGCGCGG 2868  
Oy 329 ----- 329  
Db 2869 TTACCCACAGAGTATGTGAGATTATGTGACTGTGTGAGACTGATGACTTTT 2928  
Oy 329 ----- 329  
Db 2929 GGGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCTGACTGCTGMAAAGGG 2988  
Oy 329 ----- 329  
Db 2989 GAGCGCTGCCCCAGCCCCCATCTGCACATTTGATGTCTATCATGATCATGTCAATGT 3048  
Oy 329 ----- 329  
Db 3049 TGGATGATTGACTGTGAATGTGCGCAGAAATTCGCGAGATGTGATTTCCCGC 3108  
Oy 329 ----- 329  
Db 3109 ATGGCAGGAGCCCCCAGCGCTTGTGTGTCATCCAGAAATGAGGACTTGGCCCAAGCACT 3168  
Oy 329 ----- 329  
Db 3169 CCCTTGGACAGACCTTCTACCGCTCACTGCTGGAGACGATGACATGGGAGACTGTGTG 3228  
Oy 330 ----- 330  
Db 3229 GATGCTGAGAGTATCTGTACCCAGCAGGCGCTTCTGTCCAGACCTGCCCCGGGGC 3288  
Oy 331 ALaGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyLysP 350  
Db 3289 GCTGGGCGATGTCTCCACAGCGACCGCAGCTCATCTTACCAGGAGTGGGGTGGGAC 3348





D	b	1129	AMCAGAGG	GTGACAGCAGAGAGATGGACACACAGCGGTGTGAGAACTGCACGACAGCCCTGT	1188
O	y	329	-----	-----	329
D	b	1189	GCCGAGTGTGCTATGCTGTGGCATGAGACACTTGCAGAGGTGAGGGCATTTACAGT	1248	
O	y	329	-----	-----	329
D	b	1249	GCCAAATTCCAGAGATTGCTGCTGCCTCAGAGAAAGATCTTTGGAGCCTTGACATTTCTGCCG	1308	
O	y	329	-----	-----	329
D	b	1309	GAGAGCTTTGATGGGAGACCCAGCCTCCAAACACTGCCCCCGCTCCAGACAGAGAGOTCCA	1368	
O	y	329	-----	-----	329
D	b	1369	GTTGTTGAGACTGTGAGAGAGATCACAGGTTACTATACATCTCAGATGCGCGAGAC	1428	
O	y	329	-----	-----	329
D	b	1489	GCGCCTACTCGCTGACCTCGCAAGGGCTGGGCATCAGCTGGGCTGCGCTCACTG	1548	
O	y	329	-----	-----	329
D	b	1549	AGGNACTGGGCACTGGACTGGCTTCTTCATCCACATAACCCACCTCTGCTTGTCAC	1608	
O	y	329	-----	-----	329
D	b	1609	ACGGTGCCTGGGACACAGCTCTTTGGAACCGCACCAAGTCTGCTCCACACTGCCAAC	1668	
O	y	329	-----	-----	329
D	b	1669	CGGCCAGAGCAGAGTGTGTGGCGAGGGCTGGCCTGCCACAGCTGTGCGCCGAGG	1728	
O	y	329	-----	-----	329
D	b	1729	CACGTGTGGGATCCAGGGGCCACACCACTGATGTCAACTGACGCCAGATTCTTGGGGCCAG	1788	
O	y	329	-----	-----	329
D	b	1789	GAGTGCCTGGAGAGATCCGAGTACTGACAGGGCTCCCGAGGATATGTAAATCCAG	1848	
O	y	329	-----	-----	329
D	b	1849	CACGTTTGCCGTGCCACCTGATGATGACGCCAGAAATGGTCACTGATGACCTGTTTGA	1908	
O	y	329	-----	-----	329
D	b	1909	CCGAGAGCTGACAGTGTGTGGCCTGTGCCACTATAAGACCCTCCCTTCGCGTGGC	1968	
O	y	329	-----	-----	329
D	b	1969	CGCTGCCCAAGCGTGTGAAACCTGACCTCTCTACATGCCATCTGGAATTTCCAGAT	2028	
O	y	329	-----	-----	329
D	b	2029	GAGAGGGCGCATCCAGCCTTGCCCATCACTGCACCCACTCTGTTGTGAGACTGGAT	2088	
O	y	329	-----	-----	329
D	b	2089	GACAAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTGTGACGTCAATATCTTGCGGTG	2148	
O	y	329	-----	-----	329
D	b	2149	GTTGGCATTCCTGCTGCTGCTGTGTTGGGGGTGTCTTTGGGATCCTCATCAAGACGAG	2208	
O	y	329	-----	-----	329
D	b	2209	CAGCAGAAATCCGAGATACAGATGCGGAGACTGCTCAGAAACAGGACTGGTGAG	2268	

[illegible]

```

QY 351 LeuThLeuGlLysLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 370
DB 3349 CTGACACTAGGCGCTGGAGCCCTTGAAGAGAGAGCCGCCAGCTCTCCACTGGCCCTCC
QY 371 GUGUValAGLysSerAspValAlaPheAspGluAspLeuGlyMetGlyAlaAlaLysGlyLeu 390
DB 3409 GAAAGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410
DB 3469 CAAAGCTCCCAACATGAGACCCAGCCCTTACAGCGGTACAGTGAAGAGAGCCACAGTA 3528
QY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430
DB 3529 CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 431 TyrValAsnGlnProAspValAlaArgProGlnProProSerProArgGluGlyProLeuPro 450
DB 3589 TATGTGAACCAAGCAGATGTTCCGCCAGCCCTTCCGCCAGAGAGGCGCTCTGCTGCT
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn 470
DB 3649 GCTGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 471 GlyValValLysAspValAlaPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490
DB 3709 GGGGTCTCTCAAGACGTTTTCCTTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
DB 3769 CCCCAGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 511 AsnLeuTyrTyrTyrAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530
DB 3829 AACCTATTACTGGACCAAGAGCCACCAAGAGGGGGCTCCACCCAGACCTTCAAA 3888
QY 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaPro 550
DB 3889 GGGGACCTTACGGCAGAAACCCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 551 AlaArgSerProSerPro 556
DB 3943 ACCGAGAGGCCAAGTCCG 3960

```

RESULT 6  
 US-08-625-101-1  
 ; Sequence 1, Application US/08625101  
 ; Patent No. 5869445  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE  
 ; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION  
 ; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SEED and BERRY LLP  
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue  
 ; CITY: Seattle  
 ; STATE: Washington  
 ; COUNTRY: USA  
 ; ZIP: 98104-7092  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/625,101  
 ; FILING DATE: 01-Apr-1996  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:

```

NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3768 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3765
US-08-625-101-1
Alignment Scores:
Pred. No.: 8,72e-160 Length: 3768
Score: 2405.50 Matches: 509
Percent Similarity: 41.08% Conservative: 0
Best Local Similarity: 41.08% Mismatches: 3
Query Match: 64.34% Indels: 727
Gaps: 1
US-09-821-883-2 (1-690) x US-08-625-101-1 (1-3768)
QY 35 LeuAlaArgGlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeu 54
DB 46 TTGGCCCCCGAGGCGGAGCACCACCAAGTGTCACCGGACACAGACATGAAGAGCTGGCTC 105
QY 55 ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnVal 74
DB 106 CCTGCCAGTCCCGAGACCCACCTGACATGCTCCGCCACCTCTACAGGCTGACAGAGT 165
QY 75 ValGlnGlyAsnLeuGlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGln 94
DB 166 GTGCAGGGAACCTGGAACCTACCTGCTGCCACCAATGCGACGCTGCTCTCTCAG 225
QY 95 AspIleGlnGluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValPro 114
DB 226 GATATCCAGAGAGGTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 115 LeuGlnArgLeuArgGlyLeuValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAla 134
DB 286 CTGCAGAGGCTGCGGATGTCGAGGACCCAGCCTCTTGGAGCAACTATGCTGCTGCTGCTGCTGCTGCTG
QY 135 ValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGly 154
DB 346 GTGCTAACAATGAGAGACCCGCTGACATACCAACCCCTGTCACAGGGGCTCCCAAGA 405
QY 155 GlyLeuArgGluLeuGlnLeuArgSerLeuThrGlnLeuLysGlyValLeuIle 174
DB 406 GGCCTGGGGAGCTGCGAGCTTCGAAAGCTCAGACAGATCTTGAAGAGAGGGCTGCTGATC 465
QY 175 GlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeuThrPylAspIlePheHisLys 194
DB 466 CAGCGGAACCCCGAGCTCTCTACAGACAGCAATTTTGGAGAGCAATCTTCCAAAG 525
QY 195 AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCys 214
DB 526 AACACACAGCTGGCTCTACACATGATGACACCAACCGCTCTGCGGCTGCGACCCCTGT 585
QY 215 SerProMetCysLysGlySerArgCysTyrPylGluSerSerGluAspCysGlnSerLeu 234
DB 586 TCTCCGATGTGTAAGGCTCCCGCTGCTGGGAGAGAGTTCTGAGATTTGCAAGAGCTG 645
QY 235 ThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCys 254
DB 646 ACCGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY 255 CysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys 274

```

Dh 706 TGCCATGAGCAGTGTCTGCGGCTGCACGCGGCCCAAGCACTGTGACTGCGCTGC 765  
Qy 275 LeuH1spheasnH1serglYlleCysgluLeuH1scYsProalalauValThrTyrasn 294  
Dh 766 CTCACCTTCAACCACTGATCTGTGAGCTGAGCTGCGACCTGCGTCACTTACCAAC 825  
Qy 295 ThrAspThrPhegluSerMetProasnProgluNlyArgTyrThrPheglYAlaSerCys 314  
Dh 826 ACAAGACGTTTGAGTCCATGCCCAATCCGAGGCGGATATACATTCCGCGCAGCTGT 885  
Qy 315 ValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySer 329  
Dh 886 GTGACTCCCTGCTCCCTCACTACCTTTCTACGCGAGCTGGGATCTGTCACCTGCTGC 945  
Qy 329 329  
Dh 946 CCCCCTGCACAACAGAGGTGACAGCAGAGAGTGAACACACGCGGTGTGAGAGTGCAGC 1005  
Qy 329 329  
Dh 1006 AAGCCCTGTGCGCGAGTGTGCTATGTCTGGGCATGGAGCATTGGAGAGGTGAGGGA 1065  
Qy 329 329  
Dh 1066 GTTACAGTGCATATATCAGAGAGTTGTCTGCTGCAAGAGATCTTTGGAGCTGGCA 1125  
Qy 329 329  
Dh 1126 TTTTGGCGGAGAGCTTTGATGGGACCAGCCTTCCAAACACTGCCCGGCTCCAGCCAGAG 1185  
Qy 329 329  
Dh 1186 CAGCTCAAGTGTTGAGACTCTGGAAGATCAGAGTTACATATACATCTAGCATGG 1245  
Qy 329 329  
Dh 1246 CCGAGAGCCTGCTGACTGACTCAGCGCTTCCAGAACTGCAGATATCCGGGAGCAATT 1305  
Qy 329 329  
Dh 1306 CTGCACATGCGCGCTACTGCTGCTACCTGCAAGGCTGGGCATCAGCTGCGGCTG 1365  
Qy 329 329  
Dh 1366 CGCTACTGAGGAAGTGGGAGTGAAGTGGCCCTCATCCACATTAACACCCACTCTGC 1425  
Qy 329 329  
Dh 1426 TTTCGTGACACGAGTGCCTGGGACCAGCTTTTCGGAACCCGACCAAGCTCTGCTCAC 1485  
Qy 329 329  
Dh 1486 ACTGCCAACCGGCCAGAGAGCAGAGTGTGTGGGCGAGGGCTGCGCTGCCACCACTGTGC 1545  
Qy 329 329  
Dh 1546 GCCCGAGGGCAGTGTGGGTTCAGAGGGCCACCAAGTGTGCACTGACAGCAATTCTT 1605  
Qy 329 329  
Dh 1606 CGGGGCGAGAGTGCAGAGAGATGCCAGTACTGACAGGGGCTCCCGAGGGAGTATGTG 1665  
Qy 329 329  
Dh 1666 AATGCCAGGCACTGTTTGGCGTCCACCTGAGTGTACGCCAGAAATGCTCAGTGACC 1725  
Qy 329 329  
Dh 1726 TGTTTTGGACCGGAGGCTGACAGCACTGTGTGGCTGTGCCCACTAAGAGACCCTCTTC 1785  
Qy 329 329  
Dh 1786 TGCGTGCGCCGCTGCCAGCGGTGTGAACCTGACCTCTCTACATGCCACTGTGAA 1845  
Qy 329 329  
Dh 1846 TTTCACATGAGAGAGGCGCATGCGACCTTTGCCCATCAACTGCACCCACTCTCTGTG 1905  
Qy 329 329  
Dh 1906 GACCTGATGACAAGGCTGCCCCCGCAGACAGAGCAGCCCTGTGACGTCCATCATC 1965  
Qy 329 329  
Dh 1966 TCTGCGGTGTGGCATTTGCTGTGCTGTGTGTGGGGTGTCTTTGGGATCTCATC 2025  
Qy 329 329  
Dh 2026 AAGGACGCGACAGAAAGATCCGGAAGTACACGATGCGGAGACTGTGCAGAAACGAG 2085  
Qy 329 329  
Dh 2086 CTGTGAGCGCGCTGACACCTAGCGGAGCATGCCCAACAGCGCGAGATGCGGATCTCT 2145  
Qy 329 329  
Dh 2146 AAAGAGAGGAGCTGAGAGAGTGAAGTGTGATCTGGCGCTTTTGGCAGACTTAC 2205  
Qy 329 329  
Dh 2206 AAGGCACTGTGATCCCTGATGGGAGATGTGAATAATTCAGTGGCCATCAAGTGTG 2265  
Qy 329 329  
Dh 2266 AGGAAAAACACATCCCCCAAGCCAAAGAAATCTTAGAGAAAGATAGTATGGCT 2325  
Qy 329 329  
Dh 2326 GGTGTGGCTCCCATATGTCTCCCGCTTCTGGGCACTGTGCTGACATCCAGCGTCA 2385  
Qy 329 329  
Dh 2386 CTGTGACACAGCTTATGCGCTATGCGCTTATGACCATGTCCGGAAACCGGGA 2445  
Qy 329 329  
Dh 2446 CGCTGGGCTCCAGAGCTGTGACTGATGTATGAGATTTGCCAAGGGAGTAC 2505  
Qy 329 329  
Dh 2506 CTGGAGATGTGCGGCTGTACACAGGAGCTTGGCCCTCGGAACGTGTGTCAAGAT 2565  
Qy 329 329  
Dh 2566 CCCAACATGTCAAAATTTACAGACTTGGGCTGGCTGTGTGATTTGACAGACACA 2625  
Qy 329 329  
Dh 2626 GATACATGACAGATGGGGGCAAGTGGCCATCAAGTGAATGGCTGAGTCCATTCTC 2685  
Qy 329 329  
Dh 2686 CGCGGGGTTCAACCACAGAGTGAATGTGAGATTATGTGTGACTGTGTGGAGCTG 2745  
Qy 329 329  
Dh 2746 ATGACTTTTGGGGCCAAACCTTACGATGGATCCAGCCGGGAGATCCCTGACTGCTG 2805  
Qy 329 329  
Dh 2806 GAAAGGGGAGCGGCTGCCAGCCCGCATCTGCACATTTGATGTACATGATCATG 2865  
Qy 329 329  
Dh 2866 GTCAATGTGATGATTAAGTCTGAATGTGCGCAAGATTCCGGAGTGTGTCTGA 2925



Db	466	CAGGGAAACCCCGAGCTCTGCTACAGAGACAGATTTTGTGGAGGACATCTTCCACAG	525
QY	195	AsnaaNgInLeuaLaLeuThrLeuIleAspThrAsnaArgSerArgAlaCysHisProCys	214
Db	526	AACAACAGAGCTGCTCTCACACTGATGACACCAACACCGCTCTGGGCTCGACCACTCTGT	585
QY	215	SerProMetCysLysGlySerArgCysTyrPglYglUserSerGluAspCysIserLeu	234
Db	586	TCTCCGATGTGTAAAGGGCTCCGCTGCTGGGAGAGATTCTGAGGATTTGTAGAGCCTG	645
QY	235	ThraArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCys	254
Db	646	ACGGGCACTGCTCTGTGGCGGGTGGCTGTGGCCCTGCGAAGGGCCACTGCCCACTACTGC	705
QY	255	CysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys	274
Db	706	TGCCATGAGCAGTGTGCTGCCGGCTGCACAGGGCCCCAACACACTGTACGTGCTGGCTGC	765
QY	275	LeuHisPheLeuHisSerGlyIleCysGlyLeuHisCysProAlaLeuValThrTyrAsn	294
Db	766	CTCCACTTTCACACCACTGGCATCTGTGAGCTGCACTGCCACGCTGTGATCTACCAAC	825
QY	295	ThraSPThrPheGluSerMetProAsnProGluGlyArgTyrThrPheIleLysCys	314
Db	826	ACAGACAGTGTGAGTGCATGGCCCAATCCCGAAGGGCGGATATACATGTGGCGCCAGCTGT	885
QY	315	ValThrAlaCysProTyrPheAsnTyrLeuSerThrAspValGlySer-----	329
Db	886	GTGACTGCTGTCCCTCAACTACTTTCATACGAGAGTGGATCCGACCTCGTCTGC	945
QY	329	-----	329
Db	946	CCCCGCACAACAAGAGGTGACAGCAGAGAGATGAAACACAGCGGTGTGAGAAGTGCAGC	1005
QY	329	-----	329
Db	1006	AAGCCGTGTGCCGAGTGTGCTATGCTGTGGGATGGAGACTTGTGGAAGGTGAGGGCA	1065
QY	329	-----	329
Db	1066	GTTACCAAGTCCCAATATCCAGAGATTTCGTGCTGCAGAAGATCTTTGGAGCCTGGCA	1125
QY	329	-----	329
Db	1126	TTTTGCCGGAGACTTTGATGGGAGCCAGCTCCAACTACTGCCCGCTCCAGCCAGAG	1185
QY	329	-----	329
Db	1186	CAGCTCCAAAGTTTGTGAGACTCTTGGAAGATCAGAGTTACTATATACATTCAGAGATGG	1245
QY	329	-----	329
Db	1246	CCGGACAGCTGCTGACTCAGCGTTCACAGAACCTGCAGACTAATCCGGGGAGCAATT	1305
QY	329	-----	329
Db	1306	CTGACCAATGGCGGCTACTGGCTGACCCCTGCAGAGGGCTGGGAGTCAAGTGGGCGCTG	1365
QY	329	-----	329
Db	1366	CGCTCACTGAGAGAACTGGGCACTGAGCTGACCTCATCCACCATTAACACACCACTTGC	1425
QY	329	-----	329
Db	1426	TTCTGTCACACAGGTGCCCTGGGACCAAGCTTTTCGGAAACCCGACCAAGCTCTGCTCAC	1485
QY	329	-----	329
Db	1486	ACTGCAACGGGCGCAGAGAGAGATGTGTGGGGAGAGGGCTTGGCTTGGCACCAGACTGTGC	1545
QY	329	-----	329

Db	1546	GGCCGAGGGGACTGCTGCGGGGTCCAGGGGCCACCACCACTGTGTCAACTGCAGCCAGTTCCTT	1605
Qy	329	-----	329
Db	1606	CGGGGCGAGAGATGCGTGGAGGAATCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTG	1665
Qy	329	-----	329
Db	1666	AATGCCAGGCACTGTTTGCGCTGCCACCCCTGA6TGTCAAGCCCCAGAANTGGCTCAGTGACC	1725
Qy	329	-----	329
Db	1726	TGTTTTGGACCGGAGGCTGACCAGTGTGTGGCTGTGCCACTATTAAGGACCTCCCTTC	1785
Qy	329	-----	329
Db	1786	TGCGTGCCCGGTGCCCGACGGGTGTGAACCTGACCTCTCTTACATGCCCATCTGGAAG	1845
Qy	329	-----	329
Db	1846	TTTTCCAGATGAGGAGGGCGCATGCCACCTTGCCCATCAACTGCACCCACTCTGTGTG	1905
Qy	329	-----	329
Db	1906	GACCTGGATGACAGAGGGCTGCCCGCGGACAGAGAGCCACCCCTGTGACGTCCATATC	1965
Qy	329	-----	329
Db	1966	TCTGCGGTGGTTGGCATTCGTCTGTCGTGTGTCTTG6GGGTG6TCTTTGGGATCCTCATC	2025
Qy	329	-----	329
Db	2026	AAGCGAGCGAGCAGAGAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGAG	2085
Qy	329	-----	329
Db	2086	CTGTGGAGCCGCTGACACCTAGCGGAGCGATGCCCAACGAGCGAGATGTGGATCTTG	2145
Qy	329	-----	329
Db	2146	AAAGAGACGAGCTGAGGAAGTGAAGTGTGGATCTGGGCTTTTGGCACAGCTTAC	2205
Qy	329	-----	329
Db	2206	AAGGCATCTGATTCCTGTGGGGAGAAATGTGAATAATCCAGTGCCATCAAAAGTTTG	2265
Qy	329	-----	329
Db	2266	AGGAAACACATCCCCCAAGCCAAAGAAATCTTAGACGAGCATACGTATGGCT	2325
Qy	329	-----	329
Db	2326	GGTGTGGCTCCCATATATGTCTCCGCCCTTCTGGGCATCTGCCATCTCACAGGTGCAG	2385
Qy	329	-----	329
Db	2386	CTGTGACACAGCTTATGCCCTATGGCTGCCTTATGACCATGTCCGGGAAACCGGGA	2445
Qy	329	-----	329
Db	2446	CGCTGGGCTCCAGGACCTGTGAACGTGTATGACAGATTGCCAAGGGGATGACCTAC	2505
Qy	329	-----	329
Db	2506	CTGGAGGATGTGGGCTCGTACACAGGACTTGGCCGCTCGGAACGTGTGTCAAGAGT	2565
Qy	329	-----	329
Db	2566	CCCAACCATGTCAAAATTACAGACTTGGGGCTGGCTCGCTGGCATTTGACAGACACA	2625
Qy	329	-----	329
Db	2626	GAGTACCATCAGATGGGGCAAGTGTCCCATCAAGTGATGGCGCTGAGATCCATTCTC	2685

Oy	329	-----	322
Db	2686	CGCCGGCGGCTTACACCACACAGAGTGTGTGGAGTTATGTGTACTGTGTGGAGCTG	27454
Oy	329	-----	329
Db	2746	ATGACTTTTGGGGCAACACTTACGATGGGATCCAGCCGGAGATCCCTGACTGCTG	2805
Oy	329	-----	329
Db	2806	GAAGAAGGGAGCGGCTGCCCCAGCCCCCATTTGACACATTGANTGTCTACTGATCATG	2865
Oy	329	-----	329
Db	2866	GTCAAATGTTTGATGATTTGACTCTGAATGTCGGCCAAAGATTCCGGAGTGTGTGTGAA	2922
Oy	329	-----	329
Db	2926	TTTCTCCGCATGGCCAGGAGACCCCGACGCTTTGTGTGTCATCCAGATAGAGACTTGGGC	2985
Oy	329	-----	329
Db	2986	CCAGCCAGTCCCTTGGACACACCTTCTACCGCTCACTGCTGAGAGACATGACATGGGG	3045
Oy	329	-----	329
Db	3046	GACCTGTGATGCTGAGAGATCTGTACCCACAGAGGGTCTTCTGTCCAGACCCT	3105
Oy	330	-----GlyAlaGlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGly	347
Db	3106	GCCCCGGGCGCTGGGGGCACTGTCTCCACACAGCACCAGACTCATCTACCGAGGTGGC	3165
Oy	348	GlyGlyAspLeuThrLeuGlyLeuGluProSerGlyGluGluAlaProArgSerProLeu	367
Db	3166	GGTGGGGACCTGACACTAGGAGCTGGAGACCCTCTGGAAGAGAGGCCCCGAGCTCCACATG	3225
Oy	368	AlaProSerGlyGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAla	387
Db	3226	GCACCTCCGAAAGGGGCTGGCTCGCATGTTTGTGATGTGTACCTGGGAATGGAGGACGC	3285
Oy	388	LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnInArgTyrSerGluAsp	407
Db	3286	AAGGGGCTGCAAAAGCCCTCCACACATGACCCACAGCCCTCTACAGAGGTACGTGAGGAC	3345
Oy	408	ProThrValProLeuProSerGlyThrAspGlyTyrValAlaProLeuThrTyrSerPro	427
Db	3346	CCCAACAGTACCCCTGGCTGTGAGACTGATGAGTGTACGTTGCCCTTGACCGACAGCC	3405
Oy	428	GlnProGlyTyrValAlaGlnProAspValArgProGlnProProSerProArgGlyGly	447
Db	3406	CAGCTGAAATGTGAAACACGACGATGTTGGGCCACACCCCTTCCGCCCGGAAGAGGC	3465
Oy	448	ProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluAlaArgAlaLysThrLeuSerPro	467
Db	3466	CCCTGCTGCTGCTGCCACACTGCTGGGTCACACTGTGAAAGGCCCAAGACTCTCTCCCA	3525
Oy	468	GlyLysAlaGlyValAlaLysAspValPheAlaPheGlyGlyAlaValGluAsnProGlu	487
Db	3526	GGGAGAGATGGGTGTCGTCAAAGACGTTTTGGCTTTGGGGGGTCCGTGGAGAAACCCGAG	3585
Oy	488	TyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerPro	507
Db	3586	TACTTGAACACCCCGAGGAGAGAGTGGCCCTCAGACCCACCTCTCTGCTTCAAGCCA	3645
Oy	508	AlaPheAspAsnLeuTyrTyrTTPAspGlnAspProProGluAlaGlyAlaProProSer	527
Db	3646	GCCCTTCACAAACCTCTTACTGGGACACAGAGACCCACAGACGGGGGCTCCACCCAGC	3705
Oy	528	ThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspAlaPro	546
Db	3706	ACCTTCAAAAGGACACCTTACGGGAGAGAACCCAGATACTCTGGTCTGGACCTGGCA	3762

```

RESULT 8
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5876546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEP-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 9,996-150 Length: 2385
Score: 2258.50 Matches: 490
Percent Similarity: 64.59% Conservative: 19
Best local Similarity: 62.18% Mismatches: 126
Query Match: 60.40% Indels: 154
DB: 2 Gaps: 19

US-09-821-883-2 (1-690) x US-09-146-283-3 (1-2385)
0y 35 leuAlaArGyAlaAlaSerThGlnValCysThrGlyThrAspMetLysLeuArgLeu 54
|||
Db 56 TTGCCCCCGGAGCGCGGAGCAGCCCAAGTGTGCACCGGACACAGACATGAAGCTCGGCTC 115
55 ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyrgInGlycysGlnVal 74
116 CCGTCGACATCCCGAGACCCACTGGACATGTCGCCACACGCTTACGAGGGCTCGCAGGTG 175
75 ValGlnGlyAsnLeuGluLeuThrTyrTleuProThrAsnAlaSerLeuSerPheLeuGln 94
176 GTGCGAGGAACCTGGAACCTCACCCTACCTGCCGCCACCAATGCCAGCCCTGCTCTGCGAG 235
95 AspIleGlnGluValGlnGlnGlyTyrValValLeuIleAlaHisAsnGlnValArgGlnValPro 114
236 GATATTCACGAGGAGTGCAGGGCTAGCTGCTCATCTCCCTCAACCAACGAGGAGGCTGCCA 295
Db 236 GATATTCACGAGGAGTGCAGGGCTAGCTGCTCATCTCCCTCAACCAACGAGGAGGCTGCCA 295

```

OY 115 LeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAla 134  
 |||||  
 Db 296 CTCAGAGGCGTGGATGTTGGAGGACACCACTCTTTGAGAGCAACTATGCCCGGCC 355  
 OY 135 ValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGly 154  
 |||||  
 Db 356 GTGCTGAGACATGGAGACCCGCTGAAACATACCACTCTGCACAGGGGCTCCACAGA 415  
 OY 155 GlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIleuLeuArgGlyValLeuIle 174  
 |||||  
 Db 416 GGCCTCGGAGGCTGCAGCTTCGAGCCCTCAGAGATCTTGAAGAGGGGCTTGATC 475  
 OY 175 GlnArgAspProGlnLeuGlyArgGlnAspThrIleLeuTrpLysAspIlePheHisLys 194  
 |||||  
 Db 476 CAGCGAAGCCCGCTGCTGCTACAGAGACAGCATTTTGGAGAGCATCTTCCACAG 535  
 OY 195 AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCys 214  
 |||||  
 Db 536 AACAAACAGCTGCTCTACACTGATAGACACCAACCGCTCGGGCTGCCACCCCTGT 595  
 OY 215 SerProMetCysArgGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeu 234  
 |||||  
 Db 596 TCTCCATGTGTAGAGGCTCCGCTGCTGGGAGAGATTCTGAGATTTGTCAGACCTG 655  
 OY 235 ThrArgThrValCysAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCys 254  
 |||||  
 Db 656 AGCGGACGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715  
 OY 255 CysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys 274  
 |||||  
 Db 716 TGCATAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775  
 OY 275 LeuHisPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValTrpTrpAsn 294  
 |||||  
 Db 776 CTCACATTCACACACAGTGGATGTGTGAGCTGAGCTGCCACCTGTGGTACCTTCAAC 835  
 OY 295 ThrAspThrPheGluSerMetProAsnProGluGlyArgTrpThrPheGlyAlaSerCys 314  
 |||||  
 Db 836 ACAGACACGTTGTGATGCCATGCCCAATCCGAGGCGGATATCATTTGGCGGCACCTGT 895  
 OY 315 ValThrAlaCysProTrpTrpAsnTrpLeuSerThrAspValGlySer 339  
 |||||  
 Db 896 GTGACGTGCTGTCTCCATCAACTATTTCTACGAGCTGGAGCTGCACACCTGCTGTGC 955  
 OY 339 329  
 Db 956 CCCCTGCACACCAAGGTGACAGCAGAGATGGAACACGCGGTGTGAGAACTGCAGC 1015  
 OY 330 330  
 Db 1016 AAGCCCTGTGCGCGAGTGTGCTATGCTGTG---GGCATGGAGACACTTGGAGAGTGAGG 1072  
 OY 342 SerSerThrArgSerGly 351  
 |||||  
 Db 1073 GCAGTTACCACTGCCAATATCCAGAGATTGTGCTGCTGCAAGAAAGATCTTTGGGAGCCTG 1132  
 OY 352 ThrLeuGlyLeuLupProSerGluGluGluAlaProArgSerProLeuAlaProSerGlu 371  
 |||||  
 Db 1133 GCATTT---CTGCCGAGAGCTTTGATGGGAGACCCCTCCAAACACTGCCGCTGCCAG 1189  
 OY 372 GlyAlaGlySerAspValPheAsp 385  
 |||||  
 Db 1190 CCAGACACAGTCCAAAGTGTGTGAGACTGTGGAAGAGATCAGAGTTACATATCATCTCA 1249  
 OY 386 AlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGln 402  
 |||||  
 Db 1250 GCATG---GCCGAGACAGCTTGCCTGACCTGAGCTTTCCAGAACTGCAGATATCCGGGG 1308  
 OY 403 403  
 Db 1309 ACGAATTCACACATAGGCGCTACTGCTGAGCCCTGCAGAGGCTGGCATCAGCTGCGCT 1368  
 OY 409 409  
 |||||  
 ThrValProLeuProSer 414

Db 1369 GGGGCTGCGCTACCTGAGGAACTGGGCACTGGAGCTGGCCCTCATCACCATTAAACCCA 1428  
 |||||  
 OY 415 415  
 |||||  
 Db 1429 CCTCTCTCTGTCGACACAGCTGGCCGGAGCCAGCTTTTGGAAACCCCAACCAAC--- 1485  
 OY 432 ValAsnGlnProAspValArgProGln 1485  
 |||||  
 Db 1486 ---TCTGCTCCACTGCCAACCGGCCAGAGCAGAGCTGTGGGCGAGGCTGGCTG 1542  
 OY 444 ProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLys 463  
 |||||  
 Db 1543 CCA-----CCAGCTGTGCGCGGAGGCACTGCTGGGTCACAGGCGCCAC 1587  
 OY 464 ThrLeu 473  
 |||||  
 Db 1588 CCAGTGTGCACTGCAGCAGCTGCTGCGGCGGAGGTGCGTGGAGGAATGCCAGT 1647  
 OY 474 474  
 |||||  
 Db 1648 ACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGTTGCCGCCACCTCA 1707  
 OY 483 ValGluAsnProGluTrpLeuThr 494  
 |||||  
 Db 1708 GTGTCAAGCCCAATGAGCTGACATGACCTGTTTGGACCGAGGCTGACCACTGTGGC 1767  
 OY 495 AlaAlaPro-GlnProHisPro 507  
 |||||  
 Db 1768 CTGTGCCCACTAAGAGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1827  
 OY 507 ValPheAsnLeuThrTrpTrpAspGlnAspProProGluArgGlyAla---ProP 526  
 |||||  
 Db 1828 TACCTCTCTTACATGCCATCTG---AAGTTTCCAGATGAGAGGAGGCGCATGCCAGC 1884  
 OY 526 OserThrPhe 535  
 |||||  
 Db 1885 TTGCCCCATCACTGACACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944  
 OY 535 ArgLeuAsnProGluTrpLeuGlyLeuAspValProAlaAlaAlaProAlaArgSerPro 555  
 |||||  
 Db 1945 GCAGAGAGCCAGCCCTCTG---ACGTCCCTCGAGGACCCGCGCTGCGCCAG 1995  
 OY 555 rProSerThrGlnProTrpGluHisValAsnAlaIleGlnGluAlaArgArgLeuAsn 575  
 |||||  
 Db 1996 CCCACACACAGCCCTGGAGCATGTGAATGCCATCCAGAGGCCCCGCTCTCTGAA 2055  
 OY 575 IleuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValIleSerGluMetP 595  
 |||||  
 Db 2056 CTGAGTAGAGACACTGTGCTGAGATGAATGAACAGTAGAAGTCATCTCAGAAATGTT 2115  
 OY 595 AspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuTrpLysGlnGlyLeuArg 615  
 |||||  
 Db 2116 TGAACCTCCAGAGACCCAGCTGCTTACAGACCCCGCTGAGCTGTACAAACAGGCGCTGG 2175  
 OY 615 GGIserLeuThrLysLeuLysGlyProLeuThrMetAlaSerHisTrpLysGlnI 635  
 |||||  
 Db 2176 GGGCAGCTTCACCAAGCTCAAGGGCCCTTGAGCAATGATGGCCAGCACTACCAAGCA 2235  
 OY 635 scysProProThrProGluThrSerCysAlaThrGlnIleIleThrPheGluSerPhe 655  
 |||||  
 Db 2236 CTGCTCCACACCCGGAACCTTCGTGCAACCCAGATTAATCACTTTGMAACTTCA 2295  
 OY 655 scGluAsnLeuLysAspPheLeuValIleProPheAspCysTrpGluProValGln 675  
 |||||  
 Db 2296 AGAGAACTGAAAGACTTCTGCTTCATCCCTTTGACTGCTGGAGCCACTCCAGGA 2355  
 OY 675 uGlyAlaProProProAla 682  
 |||||  
 Db 2356 GTGAGACCGGCCAGATGAGGCT 2377  
 |||||  
 RESULT 9  
 US-08-579-823A-3

```

: Sequence 3, Application US/08579823A
: Patent No. 6080409
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Composition and Method
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESS: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/579,823A
: FILING DATE: 03-DEC-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0980
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2385 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
: US-08-579-823A-3

Alignment Scores:
Pred. No.: 9.99e-150 Length: 2385
Score: 2258.50 Matches: 490
Percent Similarity: 64.598 Conservative: 19
Best Local Similarity: 62.188 Mismatches: 126
Query Match: 60.408 Indels: 154
Gaps: 19

US-09-821-883-2 (1-690) x US-08-579-823A-3 (1-2385)
OY 35 LeuAlaArgIyAlaIaIaSerThrgInValIcysThrgIyThraSPmetLysLeuArgIeu 54
Db 56 TTGGCCCCCGGAGCGCGGACCGCCAGGTGTGCACCGCGCACAGACATGAACTCGGCTC 115
OY 55 ProIaIaSerProgluThrHisLeuaspMetLeuArgHisIeuArgInGlyValIeuVal 74
Db 116 CCTGCGACGTCCGAGACCCACCTGACATGCTCGGCCACCTCTACCAAGGCTGCCAGGTG 175
OY 75 ValInGlyAsnLeuGluLeuThrTyrlauProThraSPmetLysLeuArgIeuVal 94
Db 176 GTGAGGGAACCTGGAACCTACTACTGCTCCACACATGCGACAGCTCTCTCTCTGAG 235
OY 95 AspIleGlnGluValGlnGlyTyrlauValIeuIleAlaHisAsnGlnValArgIeuValPro 114
Db 236 GATATCCAGGAGGTGACGAGGCTAGCTCATCGCTCAACCAAGATGAGGAGGTCCTCA 295
OY 115 LeuGlnArgIeuArgIleValArgIyThrgInIeuPhegluAspAsnTyrAlaLeuAla 134

```

```

Db 296 CTGACAGAGCTGCGGATTTGCGAGGACCCAGCTCTTTGAGCAACATATGCCCTGGCC 355
OY 135 ValIeuAspAsnGlyAspProLeuAsnAsnThrProValIthrgIyAlaSerProGly 154
Db 356 GTGTCAGACATGGAGACCCGCTGACAAATACACACCCCTGTACAGAGGGGCTCCACAGA 415
OY 155 GlyLeuArgIuLeuGlnLeuArgSerIeuThrgIuIleuValGlyValIeuIle 174
Db 416 GGCTGCGGAGCTGACACTTCGAAAGCTTCAGAGATCTTGAAGGGGGTCTTGTATC 475
OY 175 GlnArgAsnProgluIeuCysTyrlauSPmetLysLeuArgIeuValIeuPheHisIys 194
Db 476 CAGCGGAACCCCGACCTGCTGCTACAGACACGATTTTGTGAAGACATTTCCACAG 535
OY 195 AsnAsnGlnLeuAlaLeuThrLeuIleAspThraSPmetSerArgAlaCysHisProCys 214
Db 536 AACAAACAGCTGGCTGTACACTGTATACACCAACACCGCTCTCGGGCTGCACCCCTGT 595
OY 215 SerProMetCysLysGlySerArgCysTrpGlyIuSerSerGluAspCysGlnSerIeu 234
Db 596 TCTCCGATGTGTAAAGGCTCCCGCTGCGGAGAGACTTCTGAGGATTTGACAGCCTG 655
OY 235 ThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThraSPcys 254
Db 656 AGCGGACACTGTCTGTGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
OY 255 CysHisGlnGlnCysAlaAlaGlyCysThrgIyProLysHisSerAspCysLeuAlaCys 274
Db 716 TGCCATGAGAGAGTGTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTGCTGCTG 775
OY 275 LeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaIleuValIthTrAsn 294
Db 776 CTCACCTTCAACACAGAGGATCTGTGAGCTGACAGCTCCAGCCCTGCTACCTTACAC 835
OY 295 ThrAspThrPhegluSerMetProAsnProgluIyArgTyrlauPhegluAlaSerCys 314
Db 836 ACAGACAGCTTTGAGTGCATCCATCCCAATCCAGAGGCGGTTATCATTTGCGGCCAG 895
OY 315 ValThraIaCysProTyrlauSerThraSPmetLysLeuArgIeuValIeuValIthTrAsn 329
Db 896 GTGACTGCTGCTGCTTACAACTTCTTACGAGAGCGGATGCTGACCCCTGCTGCTGCTG 955
OY 329 ----- 329
Db 956 CCCCTGCAACAACAGAGTGTGACAGAGAGTGAACACAGCGGTGTGAGAACTGCAGC 1015
OY 330 -----GlyAlaGlyIleValIthHisArgHisArgSer 341
Db 1016 AAGCCCTGTGCGGAGCTGTCTATGCTGTG---GCCATGAGACACTTCCGAGAGGTGAGG 1072
OY 342 SerSerThraArgSerGly-----GlyAlaSPleu 351
Db 1073 GCAGTTACAGTGCATATATCCAGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1132
OY 352 ThrLeuGlyIeuGluProSerGlnGluIuAlaProArgSerProLeuAlaProSerGlu 371
Db 1133 GCATTT---CTGCGGAGAGAGCTTTGATGCGGAGCCAGCTCCCAACACTGCCCTCCAG 1189
OY 372 GlyAlaGlySerAspValIlePheAsp-----GlyAspLeuGlyIleValIeu 385
Db 1190 CCAGAGCAGCTCCAGAGTGTGTGAGACCTGTGAGAGATCACAGGTATCATATCA 1249
OY 386 AlaAlaIyGlyIeuGlnSerIeuProThraSPmetProLeuGln----- 402
Db 1250 GCATG-GCGGAGACGCTGCTGACCTGACCTGCTTCCAGAACTGCAAGTAATCCGGGG 1308
OY 403 -----ArgTyrlauSerGluAspPro----- 408
Db 1309 ACGAATTCGACAAATGCGCTTACTGCTGACCTGCAACAGGCTGGGATCAGCTGCT 1368
OY 409 -----ThraValProLeuProSer----- 414
Db 1369 GGGGCTGCGCTCACTGAGGAGACTGGGAGTGGAGCTGAGCTTCAATCCATTAACACCA 1428

```



```

Qy 415 -----GluThrAspGlyTyr-ValAlaProLeuThrCysSerProGlnProGluTyr 431
Db 1429 CCTGCTGCTGCTGCACAGCGTGGCCCTGGAGACAGCTTTTGGGAACCCGCAACAGC--- 1485
Qy 432 ValAsnGlnProAspValArgProGln-----ProProSer 443
Db 1486 ---TCTGCTCCACACTGCACACCGGCCAGAGAGAGTGTGTGGCGAGGCGCTGGCCTG 1542
Qy 444 ProArgGlnGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlnAlaLys 463
Db 1543 CCA-----CCAGCTGTGGCGCCGAGGAGCAGCTGGGGGTCCAGGCGCCAC 1587
Qy 464 ThrLeu-----SerProGlyLysAsnGlyValVal----- 473
Db 1588 CCACTGTGTCAACTGCAGCCAGCTTCTTGGGGCCAGAGAGTGGAGAGATGCCAGT 1647
Qy 474 -----LysAspValPheAlaPheGlyGlyAla 482
Db 1648 ACTGACAGGCGCTCCCGAGGAGATGTGAATGCAGAGCAGTGTGGCCGACCCCTGA 1707
Qy 483 ValGluAsnProGluTyrLeuThr-----ProGlnGlyGly 494
Db 1708 GTGTCAAGCCCGAGATGGCTCAGTACCTGTTTGGACCGGAGCGCTGACAGTGTGGC 1767
Qy 495 AlaAlaPro-GlnProHisPro-----ProProAlaPheSerPr 507
Db 1768 CTGAGCCACATATAGAGCCCTCTGCTGGTGGCGCCGCTGCCAGCGGTGGAAC 1827
Qy 507 AlaPheAspAsnLeuTyrTyrPaspGlnAspProGlnArgGlyAla---ProPr 526
Db 1828 TGACCTCTCTACATGCCCATCTGG---AGCTTTCAGATGAGGAGCGCATGCCACCC 1884
Qy 526 oSerThrPhe-----LysGlyThrProThrAl 535
Db 1885 TTGCCCATCACTGACCCACCTCTGTGTGACCTGGATGACAGGGCTGCCCGCGCA 1944
Qy 535 AsnAsnProGluTyrLeuGlyLeuAspValProAlaAlaProAlaArgSerProSe 555
Db 1945 GCAGAGAGCCAGCCCTG-----ACGTCCCTGAGGAGCCGCGCCCTGCCAG 1995
Qy 555 rProSerThrGlnProTrpGluHisValAsnAlaIleGlnAlaLarArgLeuLeuAs 575
Db 1996 CCCAGGACACAGCCCTGGGAGCATGTGAATCCATCCAGAGGCGCCGCTCTCTGAA 2055
Qy 575 nLeuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValIleSerGluMetPh 595
Db 2056 CCTGAGTAGACACACTGCTGTGAGATGAATGAACATAGACATCATCAGAAATGT 2115
Qy 595 AspLeuGlnGluProThrCysLeuGlnThrArgLeuGlnLeuTyrLysGlnGlyLeuAr 615
Db 2116 TGACCTCCAGAGGAGGAGCTCTGCTACAGACCCGCTGAGCTGACAGAGGCGCTCG 2175
Qy 615 gGlySerLeuThrLysLeuLysGlyProLeuThrMetAlaSerHisTyrLysGlnH 635
Db 2176 GGGGAGGCTTCAACCAAGCTCAAGGGGCGCTTGACCATGATGGCCAGCCACTCAAAACACA 2235
Qy 635 sCyProProThrProGluTyrThrSerCysAlaThrGlnIleLeuThrPheGluSerPheLy 655
Db 2236 CTGCCCTTCAACCCCGGAACCTTCTGTGCAACCCAGATTAACCTTTGAAAGTTTAA 2295
Qy 655 sGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpGluProValGlnGl 675
Db 2296 AGAGAACTGAAGAGCTTTGCTGTGTATCCCTTTACGTGCGGAGCCAGTCCACGA 2355
Qy 675 uGlyAlaProProProAla 682
Db 2356 GTGAGACCGGCCAGATGAGGCT 2377

```

RESULT 10  
 US-09-344-195-3  
 ; Sequence 3, Application US/09344195  
 ; Patent No. 6210662

```

GENERAL INFORMATION:
APPLICANT: Laus, Reiner
                Wu, Hongyu
                Ruegg, Curtis L.
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,195
FILING DATE: 24-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2385 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
HYPOHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3
Alignment Scores:
Pred. No.: 9,99e-150 Length: 2385
Score: 2258.50 Matches: 490
Percent Similarity: 64.59% Conservative: 19
Best Local Similarity: 62.18% Mismatches: 126
Query Match: 60.40% Indels: 154
Gaps: 19
US-09-821-883-2 (1-690) x US-09-344-195-3 (1-2385)
Qy 35 LeuAlaArgGlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeu 54
Db 56 TTGCCCGCCGAGGAGCGGAGGAGCAACCAAGTGTGACCGGCGACAGACATGAACCTGGGCTC 115
Qy 55 ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnVal 74
Db 116 CCTGCACATCCCGAGAGACCACCTGACATGCTCCGCCACCTTACCAAGGCTGCCAGGTG 175
Qy 75 ValGlnGlyAsnLeuGlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGln 94
Db 176 GTGCAGGGAACCTGGAACCTACCTACCTGCCACCAATGCCAGCGTCTCTCTCCGAG 235
Qy 95 AspIleGlnGluValGlnGlyTyr-ValLeuIleAlaHisAsnGlnValArgGlnValPro 114
Db 236 GATATCCAGGAGGTGACGGGCTACGCTCATTCGCTACACCAACGAAGTAGAGCAAGTCCCA 295

```

QY 115 LeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAla 134  
DB 296 CTGAGAGGGCTGCGGATGTTGCGAGGACCCAGCTTTTGAGGACACTATAGCCCTGGCC 355  
QY 135 ValLeuAspAsnGlyAspProLeuAsnAsnThrProValThrGlyAlaSerProGly 154  
DB 356 GTGCTTAACATGGAGACCGCTGACAAATACACCCCTGTCACAGGGGCTCCCGACGA 415  
QY 155 GlyLeuArgGlnLeuGlnLeuArgSerLeuThrGlnLeuLeuGlyValLeuIle 174  
DB 416 GGCTGGGGAGCTGCAGCTTCGAGGCTCAGAGATCTTGAAGAGAGGGCTTTATC 475  
QY 175 GlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeuTryptylsAspIlePheHisLys 194  
DB 476 CAGGGAGACCCCGAGCTCTGTACAGACACGATTTTGGAAGGACATCTTCCACAG 535  
QY 195 AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCys 214  
DB 536 AACAAACAGCTGGCTCTCACACGTGATGACACCAACCGCTCTGGGCTCCACCCCTGT 595  
QY 215 SerProMetCysLysGlySerArgCysTyrGlyLysLeuSerGluAspCysGlnSerLeu 234  
DB 596 TCTCCGATGTAGTAGGCTCCCGCTGCTGGGAGAGATTCTGAGATTGTCAAGAGCTG 655  
QY 235 ThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCys 254  
DB 656 ACGGCACTGTCTGTGCGGTGGCTGTGCGCGCTGCAAGGGGCGCCACTGACATGCTGC 715  
QY 255 CysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys 274  
DB 716 TGCATAGACAGTGTGCTGCGGCTGCGAGGCGCCCAAGCACTGTGCTGCTGCGCTGC 775  
QY 275 LeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsn 294  
DB 776 CTCGACTTCAACACACAGTGCATCTGTGAGCTGACATGCCAGCCCTGGACACTACAC 835  
QY 295 ThrAspThrPheGluSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCys 314  
DB 836 ACGAGACAGTTTGTAGTCCATAGCCCAATCCGAGGGCGGATATCATTTGCGGCGAGCTGT 895  
QY 315 ValThrIleCysProTyrAsnTyrLeuSerThrAspValGlySer 329  
DB 896 GTGACTGCTGCTCCCTCAACACTTCTTACAGGACGTGGATCTCACCCTGCTGTC 955  
QY 329 329  
DB 956 CCCCTGCACAACAGAGTGCAGAGAGATGAACACAGCGGTGTGAAGATGCAGC 1015  
QY 330 330  
DB 1016 AAGCCCTGTCGCCGAGTGTCTATGTCTG---GGCATGGAGCACTTGGCAGAGGTGAGG 1072  
QY 342 SerSerThrArgSerGly 351  
DB 1073 GCAATTACAGATGCATATCCAGAGATGTTGTCGCTGCAGAAAGATCTTTGGAGGCT 1132  
QY 352 ThrLeuGlyLeuGluProSerGlnGlnAlaProArgSerProLeuAlaProSerGlu 371  
DB 1133 GCATTT---CTGCCGAGAGACTTTGATGGGAGCCCAACCTCCAGCCGCTCAG 1189  
QY 372 GlyAlaGlySerAspValPheAsp 385  
DB 1190 CCAAGACAGCTCCAGAGTTTGAAGACTGTGAAGAGATCACAGATATCAATCATCTCA 1249  
QY 386 AlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGln 402  
DB 1250 GCATG---GCCGACAGCTGCTGACCTCAGCGTCTTCCAGAACCTGCAGATATCCGGG 1308  
QY 403 403  
DB 1309 ACGAATTCGACAAATGAGCGCTACTGCTGACCCCTCAAGGCGTGGCATCAGCTGCT 1368  
QY 409 409  
DB 414 414

DB 1369 GGGGCTGCTCACTGAGGAACCTGGGAGTGGAGCTGGCCCTCATCCACCATACACCA 1428  
QY 415 415  
DB 1429 CCTGCTGTGTCACACAGGTGCTGGACACGCTTTTGGAAACCCGACCAAG--- 1485  
QY 432 ValAsnGlnProAspValArgProGln 1485  
DB 1486 ---TCTGCTCACACCTCCCAACCGGCCAGAGAGAGAGTGTGTGGCGAGGCGCTGGCTG 1542  
QY 444 ProArgGlnGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlnAlaGlyLys 463  
DB 1543 CCA---CCAGCTGTGCGCCGAGGGCAGCTGCTGGGTCCAGGCGCCAC 1587  
QY 464 ThrLeu 473  
DB 1588 CCACTGTCACTGACGACGCCAGTTCTTCCGGGGCCAGAGAGTGCCTGAGAGATGCCAGT 1647  
QY 474 474  
DB 1648 ACTGACAGGGCTCCCGAGGAGTATGTAATGCCAGGACACTGTTGCCGTGCCACCTGA 1707  
QY 483 ValGluAsnProGluTyrLeuThr 494  
DB 1708 GTGTACAGCCCGAGATGGCTCAGTACCTGTTTGGACCGAGGCTGACAGCTGTGTCG 1767  
QY 495 AlaAlaPro---GlnProHisPro 507  
DB 1768 CTGCGCCACTATPAGAACCTCCCTTCTGCGTGGCCCGCCGACCGGTGGAAC 1827  
QY 507 AlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGluArgGlyAla---ProPr 526  
DB 1828 TGACCTCTCTACATGCCATCTG---AAGTTCCAGATGAGAGGCGCATGCCAGCC 1884  
QY 526 OserThrPhe 535  
DB 1885 TTGCCCATCAACTGCACCCACTCTGTGTGAGCTGATGACAGGGCTGCGCCGCGCA 1944  
QY 535 AGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaProAlaIleArgSerProse 555  
DB 1945 GCAGAGAGCCAGCCCTGTG---ACGTCCCTGAGAGGACCGCCGCTGCGCCAG 1995  
QY 555 rProSerThrGlnProTrpGlnHisValAsnAlaIleGlnGluAlaArgArgLeuAs 575  
DB 1996 CCCGAGCAGACAGCCCTGGGAGCATGTGAATGCCATCCAGAGGCGCGCTCTCCAGAA 2055  
QY 575 nLeuSerArgAspThrAlaIleGluMetAsnGluThrValGluValIleSerGluMetPh 595  
DB 2056 CTGAGTAGAGACACTGCTGCTGATGATGAATGAACAGTGAATCTCTCAGAAATGTT 2115  
QY 595 eAspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuTyrLysGlnLeuArg 615  
DB 2116 TGACCTCCAGAGGCGGACCTGCTGATGATGAATGAACAGTGAATCTCTCAGAAATGTT 2175  
QY 615 gGlySerLeuThrLysLeuGlyGlyProLeuThrMetCalaSerHisTyrLysGlnI 635  
DB 2176 GGGGAGGCTCCACCAAGTCAAGGGCCCTTGACCATGATGCGCAGCCACTACAAACGCA 2235  
QY 635 sCysProProThrProGluThrSerCysAlaThrGlnIleIleThrPheGluSerPhe 655  
DB 2236 CTGCGCTCCAAACCCGGAACCTCTGCTGCAACCCCAATATATCACTTTAAGATTCA 2295  
QY 655 sGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTyrProValGlnI 675  
DB 2296 AGAGAACCTGAAGAGACTTGTGCTGTCAATCCCTTTGATGCTGCGGAGCAGTCCAGGA 2355  
QY 675 uGlyAlaProProProAla 682  
DB 2356 GTGAGACCGGCCAGATGAGCT 2377  
RESULT 11  
US-08-229-515A-14

Sequence 14, Application US/0822515A  
 Patent No. 5518885  
 GENERAL INFORMATION:  
 APPLICANT: RAZIUDIN  
 APPLICANT: SARAR, FAZLUL H  
 TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN  
 TITLE OF INVENTION: NEOPLASTIC DISEASE  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: NEEDLE & ROSENBERG PC  
 STREET: 127 Peachtree Street, Suite 1200  
 City: Atlanta  
 STATE: Georgia  
 COUNTRY: usa  
 ZIP: 30303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/229,515A  
 FILING DATE: 19 APR 1994  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: PERRYMAN, DAVID G  
 REGISTRATION NUMBER: 33,438  
 REFERENCE/DOCKET NUMBER: 1414, 608  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 404-688-0770  
 TELEFAX: 404-688-9880  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3955 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-229-515A-14



QY 494 GlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 513  
 Db 3635 GGACATGCTCTCCGCCGCCACCTTCTCTGCTTCACGCCAGCTTTGACAACTCTAT 3694  
 QY 514 TyrTrpAspGlnAspProProGlnArgGlyAlaProProSerThrPheLysGlyThrPro 533  
 Db 3695 TACTGGGACCAAGACTCATCGAGCAGCGGCTCCACCAAGTAACCTTTGAAGGAGACCC 3754  
 QY 534 ThrAlaGlnAspProGlnTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSer 553  
 Db 3755 ACTGAGAGAACCTGTGACTACCTAGCCTGTGATGTATGAGACGCTGCGAGCT 3814  
 QY 554 Pro-----SerProSerThrGlnProTyrGln 562  
 Db 3815 CCTGTGCTTCAGACTGGGGAAGCGCTGACTTGTGTCTCCATCCCAACAAAGCGGAG 3874  
 QY 563 HisVal 564  
 Db 3875 AGGCTC 3880  
 RESULT 12  
 US-08-645-865-14  
 : Sequence 14, Application US/08645865  
 : Patent No. 5654406  
 : GENERAL INFORMATION:  
 : APPLICANT: RAZIUDIN  
 : APPLICANT: SARKAR, FAZLUL H  
 : TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN  
 : TITLE OF INVENTION: NEOPLASTIC DISEASE  
 : NUMBER OF SEQUENCES: 19  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: NEEDLE & ROSENBERG PC  
 : STREET: 127 Peachtree Street, Suite 1200  
 : CITY: Atlanta  
 : STATE: Georgia  
 : COUNTRY: USA  
 : ZIP: 30303  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/645,865  
 : FILING DATE: 14 MAY 1996  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: PERRYMAN, DAVID G  
 : REGISTRATION NUMBER: 33,438  
 : REFERENCE/DOCKET NUMBER: 1414,608  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 404-688-9880  
 : TELEFAX: 404-688-9880  
 : INFORMATION FOR SEQ ID NO: 14:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 3955 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : US-08-645-865-14  
 Alignment Scores:  
 Pred. No.: 4,98e-129 Length: 3955  
 Score: 1968.00 Matches: 443  
 Percent Similarity: 35.87% Conservatve: 24  
 Best Local Similarity: 34.02% Mismatches: 83  
 Query Match: 52.63% Indels: 752  
 DB: 1 Gaps: 4  
 US-09-821-883-2 (1-690) x US-08-645-865-14 (1-3955)  
 QY 3 AlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeuLeu 22

Db 11 GCGCGAATGATCATCATGAGACTGGCGGCTGTGCCCTGGGGGTTCTCTGCGCTTC 70  
 QY 23 PhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42  
 Db 71 -----CTGCCCCCGGATGCGGGGACC 94  
 QY 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62  
 Db 95 CAAGTGTATACCGGACAGACATGATGAGTGGCGCTCCCTGCCAGTCCAGAGCCACCTG 154  
 QY 63 AspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGluLeuThr 82  
 Db 155 GACATGCTCCGCCACCTGTACAGAGGCTGTGAGTAATGACAGGCAACTTGGAGCTTACC 214  
 QY 83 TyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnValGlnGlyTyr 102  
 Db 215 TACGTGCTGCCAATGCCAGCTCTCATTTCTGTCAGACATCCAGAAAGTTCCAGGTTAC 274  
 QY 103 ValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgGlyIleValArg 122  
 Db 275 ATGCTCATCGCTCACACACAGGTGAAGCGGCTGCCACAGTGCMAAGCTCGCATCTGAGA 334  
 QY 123 GlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeu 142  
 Db 335 GGGACCCAGCTCTTTGAGAGACAAATATGCCCTGTGTGCTAGACACAGAGATCTCAG 394  
 QY 143 AsnAsnThrThrProValThr---GlyAlaSerProGlyGlyLeuArgGluLeuLeu 161  
 Db 395 GACAATGTGCGCCGCTCCACCCAGGACAGAACCCAGAGGGGCTCGGAGACTGACGCTT 454  
 QY 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAspProGlnLeuCys 181  
 Db 455 CGAAGTCTCAGAGATATCTGTAAGGAGAGATTGATCCCGGGAACCTCAGCTTCGC 514  
 QY 182 TyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201  
 Db 515 TACAGAGACATGTTTGTGGAAGAGACTGTCCGCAAGAAATACCAACTGCTCTGTC 574  
 QY 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221  
 Db 575 GATATAGACACCAATCGTTCGCCGGCTGTCCACCTGTGTCGCCCGCTCGCAACAACT 634  
 QY 222 ArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThrValCysAlaGly 241  
 Db 635 CACTGTGGGAGAGATCCGGAAGACTGTGCAGATCTTGACTGGCACTCTGTACCACT 694  
 QY 242 GlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAlaAla 261  
 Db 695 GGTGTGCCCGGTGCAAGGGCGCGCTGCCACTGACTGTGSCATGAGAGAGTGCCTCGCA 754  
 QY 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281  
 Db 755 GGCTGCAGGGGCCCAACATTTCTACTGCGGCTCGCTCCACTCAATCATATGATGTGT 814  
 QY 282 IleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMet 301  
 Db 815 ATCTGTAGCTGCACTGCCAAGCTCTGTACCTTCAACACAGACACTTTTAAGTCCATG 874  
 QY 302 ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn 321  
 Db 875 CACAACCTGAGAGGTGCTACACCTTTGTGTGTCAGCTGTCGATCCACCTGCCCCCAAC 934  
 QY 322 TyrLeuSerThrAspValGlySer----- 329  
 Db 935 TACTGTCTAGGAGAGTGGATCTGACACTGTGTGTGTCCCCGAATTAACCAAGAGTTC 994  
 QY 329 ----- 329  
 Db 995 ACAGCTGAGAGCGAACAACAGCGTTGTGAGAAATGACACCAAGCCCTGTGCTGATGTGC 1054  
 QY 329 ----- 329

Db 1055 TATGCTGGGCGATGAGCACTTTCGAGGGCCAGGCCATCACCAGTGCATATGTCAG 1114  
QY 329 ----- 329  
Db 1115 GAGTTGATGGCTGCAGAAAGATCTTTGGAGCCTGGCATTTTTGGCGGAGAGCTTTGAT 1114  
QY 329 ----- 329  
Db 1175 GGGAGCCCTCTCCGCGATTGCTCCGTGAGGCGCTGAGAGCTCCAAAGTGTTCGAAACC 1234  
QY 329 ----- 329  
Db 1235 CTGGAGAGATCACAGTTACCTGTACATCTCAGCANTGCCAGACACTCTCCGTGACCTC 1294  
QY 329 ----- 329  
Db 1295 AGTGTCTCCAGAACCTTCGAATCATTCGGGAGAGATTCACAGATGGCGGCTACTCA 1354  
QY 329 ----- 329  
Db 1355 TTGACACTGCAGAGCCTGGGGATCCACTCGCTGGGCTCGCTCCTGCGGAGACTGGGC 1414  
QY 329 ----- 329  
Db 1415 AGTGGATTGGCTGTGATTCAACCGCAAGCCCATCTGTCTTTGTACACACTGTACCTTGG 1474  
QY 329 ----- 329  
Db 1475 GACCAGCTCTTCGGAACCCACATCAGGCCCTGTCCACAGTGGGAACCGGCGGAAAG 1534  
QY 329 ----- 329  
Db 1535 GACTTGTGCTCTCGAGCGGCTGTGTGTACTCACTGTGTGCCACAGGCACTGCTGG 1594  
QY 329 ----- 329  
Db 1595 GGGCAGGCGCCACCCAGTGTGTCACTGCAGTCACTTTCTTGGGGCCAGAGTGTGTG 1634  
QY 329 ----- 329  
Db 1655 GAGAGTGCAGATGATGGAAGGGCTCCCGGAGTATGAGTGAACAAGCGTGTCTG 1714  
QY 329 ----- 329  
Db 1715 CCGTGTACCCCGAGTGTCAAGCTCAAAAGACTCAGAGACTGCTTTGGATCGGAGCT 1774  
QY 329 ----- 329  
Db 1775 GATGATGTGAGGCTGCGGCCACTACAGAGACTGCTCTCTGTGTGTGCTGCTGCC 1834  
QY 329 ----- 329  
Db 1835 AGTGTGTGAACCGGAGCTCTCTACATCCCATCTGGAAGTACCGGATGAGAGAGGC 1894  
QY 329 ----- 329  
Db 1895 ATATGCCAGCGGTGCCCATCACTGACCCACTCTGTGTGATCTGATGAACGAGGC 1954  
QY 329 ----- 329  
Db 1955 TGCCCAAGACAGAGAGAGCCGCCGTGATTCATTCATTGCAACTGTAGAGGGCCTC 2014  
QY 329 ----- 329  
Db 2015 CTGCTGTCTCTGATCTTATGTTGTGTGTGATCTTATCAAAAGAGAGAGACAGAG 2074  
QY 329 ----- 329  
Db 2075 ATCCGAGATATACGATGCTAGGCTGTGACAGAAACTGATTAGTGAAGCGCTGACG 2134  
QY 329 ----- 329  
Db 2135 CCCAGGGAGCAATGCCCAACAGGCTCAGATGCGGATCTTAAAGAGAGAGAGCTAAGG 2194

QY 329 ----- 329  
Db 2195 AAGTGAAGGTGCTTGATCAGAGACTTTGGCACTGTCTACAAAGGCATCTGATCCA 2254  
QY 329 ----- 329  
Db 2255 GATGGAGAAATGTAATAATCCCGTGTATCAAGGTGTGAGAGAAACACATCTCT 2314  
QY 329 ----- 329  
Db 2315 AAAGCAACAAGAAATTTAGATGAAGCGATGTGATGCTGTGTGCTTCCTAT 2374  
QY 329 ----- 329  
Db 2375 GTGTCCGCTCTGCGGCATCTGCTGATCCACAGTACAGTGTGACACAGCTTATG 2434  
QY 329 ----- 329  
Db 2435 CCTACGGCTGCTTGTGACCATGTCCGAGAACCCGAGGTGCTTACGCTCCAGGAC 2494  
QY 329 ----- 329  
Db 2495 CTGCTCACTGTGTGTTCAGATTGCCAAGGGATGAGCTACTGAGAGTGTGCGCTT 2554  
QY 329 ----- 329  
Db 2555 GTACACAGGACTGTGCTGCCCGGATGTGCTAGTCAAGATCCCAACAGTCAAGATT 2614  
QY 329 ----- 329  
Db 2615 ACAGATTTGGGCTGTGCTGCTGCTGCTGACATGATGAGACAGAGTACATGAGTGG 2674  
QY 329 ----- 329  
Db 2675 GGCAGGTGCCCATCAATGATGATGATGATGATCTATCTCAGACCGCGTCAACCAT 2734  
QY 329 ----- 329  
Db 2735 CAGATGATGTGAGACTATGAGTACTGTGTGAGAGTATGACTTTGGGGCCAAA 2794  
QY 329 ----- 329  
Db 2795 CTTACGATGATCCAGCCCGGAGATCCCTGATTTGCTGGAGAAAGCAAGCCCTA 2854  
QY 329 ----- 329  
Db 2855 CCTCAGCTCCAATCTGCACATGATGATGATGATGATGATGATGATGATGATGAT 2914  
QY 329 ----- 329  
Db 2915 GACTGTGAATGTGCGCCGAGATTCGGGAGTGTGTGAGAAATTTTCAGATGAGCGAGG 2974  
QY 329 ----- 329  
Db 2975 GACCCCGAGCTTTGTGTGTATCCAGACAGAGACTTGGGCCCATCCAGCCCATGAGC 3034  
QY 329 ----- 329  
Db 3035 AGTACTTTCACGCTTACTGCTGGAAGATGATGATGATGATGATGATGATGATGATGAT 3094  
QY 330 ----- 333  
Db 3095 GAGTATGTGTGCCCCAGAGGATTTCTCCCGGAGCCCTACCCGACAGCTGGGAGC 3154  
QY 334 MetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThrLeu 353  
Db 3155 ACAGCCCATGAAAGGACCCGAGCTGCTCCACAGAGGTGAGGTGTGTGCTGACACTG 3214  
QY 354 GlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGlyAla 373  
Db 3215 GGCCTGAGCCTCTGGAGAAAGGGCCCCCAGATCTCCTGCTCCTCTGGAAAGGGCT 3274

QY 374 GlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaValGlyLeuGlnSerLeu 393  
DB 3275 GGCCTCCGATGTTGATGGTACCCTGGCAATGGGGGTAACCAAGGGCTCCACAGCCTC 3334  
QY 394 ProThrHisAspProSerProLeuGlnArgTyrSerGlnAspProThrValProLeuPro 413  
DB 3335 TCTCCACATGACCCCTCCCTACAGCGGTACAGCAGACGCCACATATTACTCTGCC 3394  
QY 414 SerGlnThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlnTyrValAsn 433  
DB 3395 CCCGAGACTGGTGGCTATGTTGGTCCCTGGCTGCAGCCCCCGAGCTGATGTGATAC 3454  
QY 434 GlnProAspValArgProGlnProProSerProArgGlnGlyProLeuProAlaAlaArg 453  
DB 3455 CAATCAGAGGTTCAGCCTCAGCCTCTTAAACCCAGAGAGGTCTCTGCCCTCCCTCCG 3514  
QY 454 ProAlaGlyAlaThrLeuGlnArgAlaValThrLeuSerProGlyLysAsnGlyValVal 473  
DB 3515 CCTGCTGGTCTACTAGAAAGACCCAGACTCTCTCTGCGGAGAAATGGGGTTGTC 3574  
QY 474 LysAspValPheAlaPheGlyGlyValAlaValGlnAsnProGlnTyrLeuThrProGlnGly 493  
DB 3575 AAAGACGTTTTTGGCTTGGGGGGGTGTGTGGAGAACCCTGAATACTTGAACGAGAGAA 3634  
QY 494 GlyAlaAlaProGlnProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 513  
DB 3635 GGCACCTGCTGCTCGGCCACCCCTTTCCTCCCTTACAGCCAGCCTTGGACAACTCTAT 3694  
QY 514 TyrTrpAspGlnAspProProGlnArgGlyAlaProProSerThrPheLysGlyThrPro 533  
DB 3695 TACTGGGACCCAGAACTATCCGACAGGAGGCTCCACCAACTTGAAGGAGACCCCC 3754  
QY 534 ThrAlaGlnAsnProGlnTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSer 553  
DB 3755 ACTCAGAGAACTGAGTACTAGCCTGGATGATCTGTATAGACAGCTGTCCAGAGCT 3814  
QY 554 Pro-----SerProSerThrGlnProTrpGln 562  
DB 3815 CCTGTGCTTTCAGAGTGGGAAAGCCTGACTGTGTGTCTCCATGCGCCCAAGAGAGAG 3874  
QY 563 HisVal 564  
DB 3875 AGGGTC 3880

RESULT 13  
US-08-422-108-2  
Sequence 2, Application US/08422108  
Patent No. 6015567  
GENERAL INFORMATION:  
APPLICANT: Hudziak, Robert M.  
APPLICANT: Shepard, H. Michael  
APPLICANT: Ullrich, Axel  
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,108  
FILING DATE: 14-Apr-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/355460  
FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/048346  
FILING DATE: 15-APR-1993  
PRIOR APPLICATION DATA: 07/354319  
APPLICATION NUMBER: 07/354319  
FILING DATE: 19-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 554C2D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1872 nucleotides  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
US-08-422-108-2  
Alignment Scores:  
Pred. No.: 6.44e-107 Length: 1872  
Score: 1647.00 Matches: 380  
Percent Similarity: 57.68% Conservative: 22  
Best Local Similarity: 54.52% Mismatches: 146  
Query Match: 44.05% Indels: 150  
DB: gaps: 20  
US-09-821-883-2 (1-690) x US-08-422-108-2 (1-1872)  
QY 41 SerThrGlnValCysThrGlnGlyThrAspMetLysLeuArgLeuProAlaSerProGlnThr 60  
DB 1 AGCACCCAGAGGUGUCACCGGACAGACAGACAGGUGGUGGUGGUGGUGGUGGUGGUGG 60  
QY 61 HisLeuAspMetLeuAsnGlnHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGln 80  
DB 61 CACUGGAGACAGUGUGCGCCGACACUUCUACAGGUGUGCCAGUGUGUGGAGAAACUUGGAA 120  
QY 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100  
DB 121 CUCACUACUUGCCACCAUUGCCAGCCUUGCCUUGCCAGGUAUACAGAGGUGUCAG 180  
QY 101 GlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120  
DB 181 GGCUGAGUGUCUACGUCUACACCAAGUGAGGAGGUGCCACUGCAGAGGUGGAGU 240  
QY 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140  
DB 241 GUGGAGAGGACACCGUCUUGGAGACACUAGUCCUGGCGUGGUGGUGGUGGUGGUGG 300  
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeuGln 160  
DB 301 CCGCUGAACAAUACCCACCCUGUCACAGGGGCCUCCCGACAGAGCCUGCGGAGCUCAG 360  
QY 161 LeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180  
DB 361 CUUGGAAAGCCUACAGAGAUUCUUGAAGAGGAGGUGUUGAUGACCGAGAACCCCGACUC 420  
QY 181 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnAlaLeu 200  
DB 421 UGCUACAGGACAGACAGAUUUUGGAGAGACAUUCUCCAAAGAACACACUGGUCUC 480  
QY 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysLysGly 220  
DB 481 ACACUGAUGAGACCAACCGCUCUGCGGCCACCCCGUUCUCCGAGUGUGAAGGCG 540  
QY 221 SerArgCysTrpGlyGlySerSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240  
DB 541 UCCCGUGUGGAGAGAGAUUCUGAGAGAUUGUGUAGACCCUGACCGCACUGUCUGGCC 600  
QY 241 GlyLysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 260





LENGTH: 1872 nucleotides  
 TYPE: Nucleic Acid  
 STRANDEDNESS: Single  
 TOPOLOGY: Linear  
 US-08-422-734-2

Alignment Scores:  
 Pred. No.: 6,44e-107 Length: 1872  
 Score: 1647.00 Matches: 380  
 Percent Similarity: 57.68% Conservatave: 22  
 Best Local Similarity: 54.52% Mismatches: 146  
 Query Match: 44.05% Indels: 150  
 DB: 4 Gaps: 20

US-09-821-883-2 (1-690) x US-08-422-734-2 (1-1872)

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60  
 Db 1 AGCACCCAGAGUGGACCGGACAGACAGAAAGCGCGCCGCCGACAGUCCCGACAGACC 60  
 Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyValAsnLeuGlu 80  
 Db 61 CACCGGACAGUGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 120  
 Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlyValGln 100  
 Db 121 CUCACGACUCCGCGCCACCAAGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
 Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120  
 Db 181 GCGUACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
 Qy 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlnAsp 140  
 Db 241 GUGGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 300  
 Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlyLeuGln 160  
 Db 301 CCGGUGAACAAUACCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360  
 Qy 161 LeuArgSerLeuThrGlnIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180  
 Db 361 CUUCCAAACCCUCCAGAGAGUUCUGAAAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
 Qy 181 CysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200  
 Db 421 UCGUACGACGACAGAUUUUGGAGAGACAGACACUCCACAGAACACACAGGCGGCGG 480  
 Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220  
 Db 481 ACACUGAUAAGACACCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
 Qy 221 SerArgCysTrpGlyGlnSerSerLysAspCysGlnSerLeuThrArgThrValLysAla 240  
 Db 541 UCCCGGCGGCGGAGAGAGUUCUGAGAGUUCUGAGAGCGGCGGCGGCGGCGGCGGCGG 600  
 Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260  
 Db 601 GUGGCGUUGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660  
 Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280  
 Db 661 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720  
 Qy 281 GlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlnSer 300  
 Db 721 GGCACUUGGAGACUCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780  
 Qy 301 MetProAsnProGlnGlyArgTyrThrPheGlnAlaSerCysValThrAlaCysProTyr 320  
 Db 781 AUGCCCAUCCGAGGCGGCGGAGUACAUUCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840  
 Qy 321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHisArgHisArg 340

Db 841 AACUACCUUUUACGAGGAGUGGAGUCCUGACAC-----CUGGUGGCGGCGGCGGCGG 894  
 Qy 341 SerSerSerThrArgSerGlyGlyLysP----- 350  
 Db 895 CAAGAGUGGACGACGAGAGAGUAGACACAGCGGUGGAGAGAGUAGAGAGAGCGGCGG 954  
 Qy 351 -----LeuThrLeuGlyLeuGlnProSerGlnGlnGlnValAlaProArgSerProLeu 367  
 Db 955 CGAGUGGCGUAGUGGUGGCGGCGGAGAGACUCCAGAGAGAGAGAGAGAGAGAGAGAG 1014  
 Qy 368 AlaProSerGlnGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAla 387  
 Db 1015 AAUUAUCCAGAGAGUUGGUGGCGGCGGAGAGAGUCCUUGGAGGCGGCGGAGUU----- 1065  
 Qy 388 LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerLysAsp 407  
 Db 1065 ----- 1065  
 Qy 408 ProThrValProLeuProSerGlnThrAspGlyTyrValAlaProLeuThrCysSerPro 427  
 Db 1066 -----CUGCGGAGAGCGUUGAGUGGAGACCCAGCGGCGGCGGCGGCGGCGG 1113  
 Qy 428 GlnProGln-----TyrValAsnGln 434  
 Db 1114 CAGCCAGAGACAGCGGCGGAGAGUUGGAGACUCCUGAAGAGAGUACAGAGUACCUAUA-CAU 1172  
 Qy 435 ProAspValArgProGlnProProSerProAlaGlyGlyProLeuProAlaAlaArgPro 454  
 Db 1173 CUCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1232  
 Qy 455 AlaGlyAlaThrLeuGlnArgAlaLysThrLeuSerProGlyLysAsnGlyValValLys 474  
 Db 1233 GGGAGAGAUUCCGACAAUAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1283  
 Qy 475 AspValPheAlaPheGlyGlyAlaValGlnAsnProGluTyrLeuThrProGlnGlyGly 494  
 Db 1284 -----CAUACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1322  
 Qy 495 AlaAlaProGlnProHisProPro----- 502  
 Db 1323 CAGUGGACUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1382  
 Qy 503 ---ProAlaPheSer---ProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro 520  
 Db 1383 GGACGAGCGUUCUGGAGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1433  
 Qy 521 GlnArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGlnAsnProGlu--- 539  
 Db 1434 GCCAGAGGACGAGUGUGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1493  
 Qy 540 TyrLeuGly-----LeuAspValProAlaAlaAlaPro 550  
 Db 1494 CUGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1553  
 Qy 551 AlaArgSerProSerProSerThrGln-----ProTrpGlnHisValAsnAlaIleGln 568  
 Db 1554 GUGCGUGGAGAGAUUCCGAGUACUGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1613  
 Qy 568 nGlnAlaArgArgLeuLeuAsnLeuSerArgAspThrAlaAlaGlnMetAsnGlnThrVal 588  
 Db 1614 C----- 1614  
 Qy 588 LgluValIleSerGlnMetPheAspLeuGlnLupProThrCysLeuGlnThrArgLeuGln 608  
 Db 1615 -----UGUUGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1634  
 Qy 608 uLeuTyrLysGlnGlyLeuArgGlySerLeuThrLysLeuLysGlyPro-----Le 625  
 Db 1635 GUGUACGCGGCGG-----AANCGCGUACAGUACUGUUUU---GGACCGGAGCGGAGACA 1685  
 Qy 625 uThrMetAlaSerHisTyrLysGlnHisCysProProThrProGluThrSerCysAl 645

Db 1686 GUGUGUGCCUGCCACUUAAGAC-----CCUCCUUCUGCCUGCCGCCUCC 1739  
 Qy 645 arthGlllellethrphegluserpelygluasnleuylasphleuLeuVal11 665  
 Db 1740 CAGC-----GUGUGAAACCUAGCCUCCUCCUACAU 1769  
 Qy 665 eProPheaspCystrpGlu---ProValGlnGlnGlyAlaLrProPro 680  
 Db 1770 GCCCAUC-----UGAAGUUCGAGUAGAGGCGCGCAUGCCAGCCU 1812  
 RESULT 15  
 US-09-146-283-1  
 ; Sequence 1, Application US/09146283  
 ; Patent No. 5976546  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Ruegg, Curtis L.  
 ; APPLICANT: Wu, Hongyu  
 ; TITLE OF INVENTION: Immunostimulatory Compositions  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dehlinger & Associates  
 ; STREET: 350 Cambridge Ave. Suite 250  
 ; City: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94306  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/146,283  
 ; FILING DATE: 03-SEPT-1998  
 ; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Judge, Linda R.  
 ; REGISTRATION NUMBER: 42,702  
 ; REFERENCE/DOCKET NUMBER: 7636-0010.21  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-324-0880  
 ; TELEFAX: 650-324-0960  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1588 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHEICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapiens  
 ; INDIVIDUAL ISOLATE: Prostatic acid phosphatase-GM-CSF  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: prostate carcinoma Lncap.FGC; PBMC  
 ; US-09-146-283-1  
 Alignment Scores:  
 Pred. No.: 2 05e-45 Length: 1588  
 Score: 770.00 Matches: 257  
 Percent Similarity: 42.14% Conservative: 54  
 Best Local Similarity: 34.82% Mismatches: 159  
 Query Match: 20.59% Indels: 268  
 DB: 2 Gaps: 28  
 US-09-821-883-2 (1-690) x US-09-146-283-1 (1-1588)  
 Qy 1 MetArGAlaAlaProLeuLeuLeuAlaAlaAlaSerLeuSerLeuGlyPheLeuPhe 20  
 Db 15 ATGAGAGCTGACCCCTCTCTGCGCAGGCAAGCAAGCCTTACCTTGGCTTCTTT 74

Qy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGlnLeuAlaArgGlyAlaAla 40  
 Db 75 CTGCTTTTCTGCTAGACCGAAGTACTATACCAAGAGTTG----- 119  
 Qy 41 SerThrGlnValCysThrGlnGlyThrAspMetLysLeuArgLeuProLaserProGlnThr 60  
 Db 119 ----- 119  
 Qy 61 HisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlyAsnLeuGln 80  
 Db 119 ----- 119  
 Qy 81 LeuThrTyrlleuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100  
 Db 119 ----- 119  
 Qy 101 GlyTyrlValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120  
 Db 119 ----- 119  
 Qy 121 ValArgGlyThrGlnLeuPheGlnuAspAsnTyrlAlaLeuAlaValLeuAspAsnGlyAsp 140  
 Db 120 -----AAGTTGTGACTTTGGTGTTCGCGATGAGAC 152  
 Qy 141 -----ProLeuAsnAsn-----ThrThrProValThrGlyAlaSer---ProGlyGly 155  
 Db 153 CGAAGTCCCATTCGACCTTTCCCATGACCCCAATAAGAAATCCATGCGCCACAAGA 212  
 Qy 156 LeuArgGlnLeuGlnLeuAspArgSerLeuThrGlnLeuLysGlnGlyValLeuIleGln 175  
 Db 213 TTTGGCCAACTCCACCCAGCTGGGCATGAGCATTAATGAACCTTGAGACTATATAGA 272  
 Qy 176 ArgAsnProGlnLeuCysTyrglnAspThrIleLeuThrLysAspIlePheHisLysAsn 195  
 Db 273 AAGAGATATAGAAATATCTGAATGAGTCC-----TATAAATCAT 311  
 Qy 196 AsnGlnLeuAlaLeuThrIleAspThrAsnArgSer---ArgAlaCysHis-----P 213  
 Db 312 GAAACAGTTATATATTCGACACACAGACGTGACCGGACTTGATGAGCTATATACAAC 371  
 Qy 213 rocYserPrometCysLysGlySerArgCysTrpGlyLysSerSerGlnuAspCysGlns 233  
 Db 372 CTGCGACCCCTGTT-----TCCCCCAAGAGGTGTACG 404  
 Qy 233 erLeuThrArgThrValCysAlaGlnGlyCysAlaArgCysLysGlyProLeuProThra 253  
 Db 405 ATCTGAA-----TCCATATCTACTC 425  
 Qy 253 spCysCysHisGlnGlnCysAlaAlaGlnCys-ThrGlyProLysHisSerAspCysLeu 272  
 Db 426 TGGCAGCCCAT-----CCCGGTGCACACAGTCTCTTCTGAAAGACAGTGG 473  
 Qy 273 AlacysLeuHisPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThr 292  
 Db 474 CTATACCTGCTTTCAGG-----AACTCCCTCGTTTCAAGAA 512  
 Qy 293 TyrAsnThrAspThrPheGlnuSerMetProAsnProGlnuArgTyrlThrPheGlyAla 312  
 Db 513 CTTCAGAGTGAAGATTGAATCAGAGAAATCCAGAAAGAGCTGAC----- 560  
 Qy 313 SerCysValThrAlaCysProTyrlAsnTyrlLeuSerThrAspValGlySerGlyAlaGly 332  
 Db 561 -----CCTTATAGAGATTTATACCTACCTTGGGAAACTTTCGA 602  
 Qy 333 GlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyLysLeuThr 352  
 Db 603 -----TTACAT-----GGCCAGAGACTTTT 623  
 Qy 353 LeuGlyLeuGlnuProSerGlnuGlnuAlaProArgSerProLeuAlaProSerGlnuGly 372  
 Db 624 -----GGA 626

Search completed: May 8, 2003, 14:53:52  
Job time : 154 secs

Oy	373	AlaGlySerAspValPheAspGlyAspLeuGlyMeGlyAlaValAlaGlyGly-----Leu	390
Db	627	ATTGGGAGTAAAGCTACGCAC---CCTTATATATTGGAGAGCTGTACAAATTTCACATTTA	683
Oy	391	GlnSerLeuProThrHisAspPro---SerProLeuGlnArgTyrSerGlu--AspPro	409
Db	684	CCCTCCCTGGGCCACTGAGACACCATGACTAAGTTGAGAGAAATGTGTCAGAATTGCGCTC	743
Oy	409	hrValProLeu---ProSerGluThrAspGlyTyrVal---AlaProLeuThrCys-----	425
Db	744	CTGCTCCCTCATGGCAATTCCACAGCAGAAAGAAATTCAGTCCCTCAAGGGGGGTCTCTG	803
Oy	426	-----SerProGlnProGluTyrValAsnGlnProAspValArgProGlnProPro-	442
Db	804	GTCAAATGAATTCCTCATCATCAGTAGAGAGACAGCACTCAAGATACCAAGCTACAAAATT	863
Oy	443	-----SerProArgGluGlyProLeuProAlaAlaArgProAlaG	456
Db	864	ATCATGTATTCTGGCATGACACTACTGTGAGGGCTGTACAGATGGCTCTAGA-----	916
Oy	456	LysAlaThrLeuGlnArgAlaValSerThr-LeuSerProGlyLysAsnGlyValValLysAsp	475
Db	917	-----TGTTTACAAAGS-----ACTCCTTCCCTCCCTAATGCTTCTTGGCAGCTTGACGGA	965
Oy	476	ValPheAlaPheGlyGlyAlaValAlaGluAsnProGluTyrLeuThrProGlnGlyValAla	495
Db	966	TTGTATC-----TTTGAGAAAGGGGAGATC-----	989
Oy	496	AlaProGlnProHisProProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrp	515
Db	990	-----TTGTGAGATGTACTATCGG 1010	
Oy	516	AspGlnAspProProGlu-----ArgGlyAlaProProSerThrPhe	529
Db	1011	AATGAGACGCAGCAGCGCGCTATCCCGTCATGCTACTCGGTGCGAGCGCTTAGC-----	1064
Oy	530	LysGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValPro-----	546
Db	1065	-----TGTCCTGTGAGAGGTTTGCTGAGAGCTGGTGGCCGTGATCCCTCAAGCTGG 1118	
Oy	547	-----AlaAla 548	
Db	1119	TCCACGGAGTGTATGACCACAAACAGCATCAAGTACTGAGAGACAGTACAGATGATCC 1178	
Oy	549	AlaProAlaArgSerProSerProSerThrGlnProTyrGlnHisValAsnAlaIleGln	568
Db	1179	GCACCGCGCGCGCTGCCGCCACCCCGACACACACCCCTGGGACATGTGATCCATCCACG 1238	
Oy	569	GluAlaArgArgLeuAsnLeuSerArgAspThrAlaAlaGlnMetAsnGluThrVal	588
Db	1239	GAGCGCGCGGCTCTCTGAACCTGAGTAGAGACACGTGCTGAGATGAATTAACACATA 1298	
Oy	589	GluValIleSerGluMetPheAspLeuGlnGluProThrCysLeuGlnThrArgLeuGlu	608
Db	1299	GAACTCATCTCAGAAATGTTTGAACCTCAGAGACCGACCTGCGCTACAGACCGCGCTGAG 1358	
Oy	609	LeuTyrLysGlnGlyLeuArgGlySerLeuThrLysLeuLysGlyProLeuThrMetMet	628
Db	1359	CTGTACAAGCAGGGCGCTGGGGGCGAGCTCTCACCAACTCAAGGGCCCTTGACATGTAGT 1418	
Oy	629	AlaSerHisTyrLysGlnHisCysProProThrProGluThrSerCysAlaThrGlnIle	648
Db	1419	GCCAGCCACTTCAAAAGCAACGTGCGCTCCACACCCGGAAACTCTCTGTGCAACCCAGATT 1478	
Oy	649	IleThrPheGluSerPheLysGluAsnLeuLysAspPheLeuValIleProDheAsp	668
Db	1479	ATCACCTTTGAAATTTTCAAAAGAAACCTGAAGACTTTCTGCTGTCTATCCCTTTTAC 1538	
Oy	669	CysTrpGlnProValGlnGluGlyAlaProProProAla 682	
Db	1539	TGCTGGGAGCGAGTCCAGAGGTAGACCGGCGCAGATGAGCGT 1580	

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 8, 2003, 12:30:24 ; Search time 2116 Seconds

(without alignments) 5281.137 million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARASLSLGLF.....EPVOEGAPPPAAHHHHHH 690

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ p2n model -DEV-rlp  
-O=/cgn2\_1/USPRO.spool/US09821883/unat\_28042003.104749.14578/app-query.fasta.1.839  
-DB-EST -QFMT-fastcap -SUFF1-p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-Biosum62 -TRANS-human40.cdi -LIST=45  
-DOCLALIGN=200 -THRM-Score-pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT-pto -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09821883.ecgn.1.1.2814.8runat.28042003.104749.14578 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGESOURCY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMOUT=120  
-WARN\_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estmu: \*  
5: em\_estlov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hlc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hlc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_rtd: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1288	34.4	1016	13	BM562913
2	1176.5	31.5	1129	14	BM802792
3	1163	31.1	4715	11	AF318349
4	1087	29.1	653	14	BM721340
5	1046	28.0	569	9	AL701765
6	976	26.1	943	14	BO958632
7	924	24.7	813	13	BI557797
8	915	24.5	614	10	BM370693
9	912	24.4	791	13	BI154872
10	901	24.1	491	14	BM790293
11	889.5	23.8	998	13	BI649877
12	881	23.6	685	9	AU123871
13	868	23.2	609	10	AW701942
14	864.5	23.1	621	12	BC283493
15	853	22.8	887	14	BO717097
16	848	22.7	786	13	BI155788
17	835.5	22.3	649	9	AA496412
18	817	21.9	552	12	BG277542
19	811	21.7	670	9	AI906012
20	808	21.6	876	14	BO769889
21	802	21.4	514	12	BF998814
22	792.5	21.2	678	13	BI555157
23	788.5	21.1	499	12	BF746234
24	788.5	21.1	499	12	BF746304
25	781.5	20.9	657	10	BM620332
26	778.5	20.8	632	10	BM655484
27	774.5	20.7	767	12	BF964111
28	774	20.7	474	9	AL701569
29	747.5	20.0	486	9	AI906364
30	747	20.0	462	12	BF834727
31	735	19.7	956	9	AL552065
32	729	19.5	508	10	BE065758
33	722.5	19.3	923	14	BO888953
34	720	19.3	450	9	AI906050
35	719.5	19.2	2456	11	AK004911
36	719.5	19.2	2662	11	AK004883
37	719.5	19.2	2936	11	AK004944
38	717	19.2	537	14	BM793744
39	697	18.6	457	10	AM372023
40	690	18.5	729	12	BG174714
41	675	18.1	588	10	AM207707
42	675	18.1	660	10	BE218982
43	675	18.1	666	9	AI912784
44	675	18.1	672	10	BE671554
45	675	18.1	695	10	BE669962

## ALIGNMENTS

RESULT 1  
LOCUS BM562913 1016 bp mRNA EST 20-FEB-2002  
DEFINITION AGENCOURT\_6566679 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5736771  
ACCESSION BM562913  
VERSION BM562913.1 GI:18809393  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1016)  
NIH-MGC http://mgs.ncl.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: c9abs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM12745 row: n column: 04  
High quality sequence stop: 637.  
Location/Qualifiers

FEATURES  
SOURCE

1. 1016  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5736771"  
/clone\_1lb="NIH\_MGC\_88"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: small intestine; Vector: PCMV-SPORT6;  
Site.1: NotI; Site.2: SalI; Cloned unidirectionally;  
oligo-dt primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

BASE COUNT 218 a 322 c 278 g 198 t

ORIGIN

Alignment Scores:

Pred. No.: 3,41e-92 Length: 1016  
Score: 1288.00 Matches: 253  
Percent Similarity: 84.03% Conservative: 10  
Best Local Similarity: 80.83% Mismatches: 23  
Query Match: 34.45% Indels: 28  
DB: 13 Gaps: 6

US-09-821-883-2 (1-690) x BM562913 (1-1016)

QY 93 LeuGlnAspIleGlnGluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGln 112  
Db 72 TTACAGCATATCCAGAGGTGTCAGGCGCTGCTCATCTGCTCAACCAAGTGGAGCAG 131  
QY 113 ValProLeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGlnAspAsnTyrAla 132  
Db 132 GTCCCACTGCGAGAGGTGCGGATGTCGAGGACCCAGCTTTGAGCAACTATGCC 191  
QY 133 LeuAlaValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSer 152  
Db 192 CTGGCCGCTGAGACATGAGACCCGCTGAAACAATACCAACCCCTGTCACAGGGGCTCC 251  
QY 153 ProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluLeuLysGlyVal 172  
Db 252 CCAGAGAGCCCTGCGGAGCTGACCTCGAAGCCTCAAGATCTTGAAGAGAGGGGTC 311  
QY 173 LeuIleGlnArgAsnProGlnLeuGlyTyrGlnAspThrIleLeuTyrPlyAspIlePhe 192  
Db 312 TTGATCCAGCGGAACCCCGCTGCTGACAGACAGATTTGTGAGAGACATCTTC 371  
QY 193 HisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHis 212  
Db 372 CACAAGAAACACACAGCTGCTCACAACATGATACACACCAACCCCTCTCGGCGCTCCAC 431  
QY 213 ProCysSerProMetCysLysGlySerArgCysTrrpGlyGlySerGluAspCysGln 232  
Db 432 CCCGTGCTCTCGATGTGAAGGCTCCCGCTGCGGAGAGAGAGTCTGAGAGATTGTACAG 491  
QY 233 SerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThr 252  
Db 492 AGCTTGAGGCGCAGCTGTCTGCGGCTGCTGCTCCCGCTGCAAGGGGCCACTGCCACT 551  
QY 253 AspCysCysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeu 272  
Db 552 GACTGCTGCGCATGAGAGTGTGCTGCGGCTGCGAGGGGCCCAAGACACTGCTGACTGCTG 611

QY 273 AlaCysLeuHisPheAsnHisSerGlyTleCysGluLeuHisCysProAlaLeuValThr 292  
Db 612 GCTGCGCTCCACTTCACACACACAGTGGCATCTGTGAGGTGACAGCCGCTGGTACC 671  
QY 293 TyrAsnThrAspThrPheGlnSerMetProAsnProGluGlyArgTyrThrPheGlyAla 312  
Db 672 TACAACACAGACAGGTGTGATGCTCATGCGCCCAATCCCGAGGCGCGTATACATTTGGCGCC 731  
QY 313 SerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerGlyAlaGly 332  
Db 732 AGCTGTGAGCTGCTGCTGCTTACAC-TACCTTTCTACGACGCGGATCTGACAC- 787  
QY 333 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 352  
Db 788 -CTGCTGCTCCCGCTCCACACAA-----GAGGTGACC 820  
QY 353 Leu-GlyLeuGluProSerGlu---GluGluAlaProArgSerProLeuAlaProSerGlu 371  
Db 821 GCAGAGATGAGAACCCACAGCGGTGTGAAAAGTGCACACACCCCTGTGCCCA----- 875  
QY 371 uGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu- 390  
Db 876 -----GATGGCTAATGGGCTGGGCGCATGAAACACTTGGC 913  
QY 391 -----GlnSerLeuProThrHisAspPro 398  
Db 914 AAAAGTGAAGGCGCAATTACCC---CATGGCCCA 947

RESULT 2  
BM802792 1129 bp mRNA linear EST 05-MAR-2002  
LOCUS AGENCOURT\_6459871 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5575087  
DEFINITION 5', mRNA sequence.  
ACCESSION BM802792  
VERSION BM802792.1 GI:19119615  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 1129)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9abs-remail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLM12324 row: m column: 08  
High quality sequence stop: 642.  
Location/Qualifiers

FEATURES  
SOURCE

1. 1129  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5575087"  
/clone\_1lb="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: PCMV-SPORT6; Site.1: NotI;  
Site.2: SalI; Cloned unidirectionally; oligo-dt primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 243 a 367 c 325 g 194 t

ORIGIN

Alignment Scores:

Pred. No.: 2.78e-83 Length: 1129  
 Score: 1176.50 Matches: 223  
 Percent Similarity: 92.24% Conservative: 3  
 Best Local Similarity: 91.02% Mismatches: 9  
 Query Match: 31.47% Indels: 10  
 DB: 14 Gaps: 2

US-09-821-883-2 (1-690) x BM802792 (1-1129)

OY 318 CysProTyrAsnTyrLeuSerThrAspValGlySerGlyAlaGlyIleValHisHis 337  
 DB 51 TGTCCA-----GACCTGCCCCGGGGCGGCGCATGGTCCACAC 92

OY 338 ArgHisArgSerSerThrArgSerGlyGlyIleAspLeuThrLeuGlyLeuGluPro 357  
 DB 93 AGCACCCAGCTCATCTACAGAGAGTGGCGTGGGAGACTACACTAGGCTGCAGCCCC 152

OY 358 SerGluGluGluAlaProArgSerProLeuAlaProSerGlyAlaGlySerAspVal 377  
 DB 153 TGTGAAG 212

OY 378 PheAspGlyAspLeuGlyMetGlyAlaAlaGlyIleGlnSerLeuProThrHisAsp 397  
 DB 213 TTTGATGGAGCTGGGGAATGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 272

OY 398 ProSerProLeuGluArgTyrSerGluAspProThrValProLeuProSerGluThrAsp 417  
 DB 273 CCCACCCCTCTACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 332

OY 418 GlyTyrValAlaProLeuThrCysSerProGluProGluTyrValAsnGlnProAspVal 437  
 DB 333 GGCTACGTTGGCCCCCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 392

OY 438 ArgProGluProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAla 457  
 DB 393 CGGCCCCAG 452

OY 458 ThrLeuGluArgAlaAlaThrLeuSerProGlyIleValAsnGlyValAlaAspValPhe 477  
 DB 453 ACTCTGGAG 512

OY 478 AlaPheGlyGlyAlaValGlnAsnProGluTyrLeuThrProGluGlyIleValAlaPro 497  
 DB 513 GCCTTTGGGGGTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 572

OY 498 GlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGln 517  
 DB 573 CAGCCCCAGCTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 632

OY 518 AspProProGluArgGlyAlaProProSerThrPheIleGlyThrProThrAlaGluAsn 537  
 DB 633 GACCCACAG 692

OY 538 ProGluTyrLeuGlyIleAspValProAla-----AlaAlaProAlaArgSerPr 554  
 DB 693 CCAGAGTACCTGGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752

OY 554 OserProSerThr 558  
 DB 753 TGAATGTCTCTCA 765

RESULT 3  
 AF318349 4715 bp mRNA linear HTC 01-JAN-2002  
 LOCUS AF318349  
 DEFINITION Homo sapiens pp3659 mRNA, complete cds.  
 ACCESSION AF318349  
 VERSION AF318349.1 GI:18027789  
 KEYWORDS HTC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 4715)

AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,  
 Man, D.F. and Gu, J.R.  
 TITLE Novel human cDNA clones with function of inhibiting cancer cell  
 growth  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 4715)  
 AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,  
 Man, D.F. and Gu, J.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-NOV-2000) National Laboratory For Oncogenes & Related  
 Genes, Shanghai Cancer Institute, 25/ln 2200 Xie-Tu Road, Shanghai  
 200032, P. R. China

FEATURES  
 source Location/Qualifiers  
 1..4715  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 gene 1..4715  
 /gene="pp3659"  
 CDS 2950..4188  
 /gene="pp3659"  
 /codon\_start=1  
 /product="unknown"  
 /protein\_id="AA15586.1"  
 /db\_xref="GI:18027790"  
 /translation="MALESILRRRTTHQSDVMSYGVYVWELMPFANPDGIPAREIP  
 DLEKGERLPDPICTIDVIMYKCMWIDSECPRELFSESRMARDPORFVIG  
 NEDGAPASPLDSTFYSILEDMDGLVDAAEYLVPOGFPCCPAPAGAGGVHRR  
 SSSTRSGGDTTLEBSREPARSPPLASGASDVPDGLGCAKGLDSELTTHP  
 SFLDRYSDPTVPLPSERDGVAPLTSPQPEYVNOQVRRPSPREGLPAPRPG  
 ATELRPLTPGKGVKVDVAFGAENPEYLVPOGALSPYLLPSPAPSTSTT  
 GTRTHSGGLHAPSKGHLRQRTOSTWMTQCCEPQGVRRSPVSSGSRGLTSAGI  
 KRWGCPPTTSRGTHARNTLS"

BASE COUNT 969 a 1446 c 1347 g 953 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2e-81 Length: 4715  
 Score: 1163.00 Matches: 222  
 Percent Similarity: 92.89% Conservative: 0  
 Best Local Similarity: 92.89% Mismatches: 9  
 Query Match: 31.10% Indels: 9  
 DB: 11 Gaps: 2

US-09-821-883-2 (1-690) x AF318349 (1-4715)

OY 318 CysProTyrAsnTyrLeuSerThrAspValGlySerGlyAlaGlyIleValHisHis 337  
 DB 3379 TGTCCA-----GACCTGCCCCGGGGCGGCGCATGGTCCACAC 3420

OY 338 ArgHisArgSerSerThrArgSerGlyGlyIleAspLeuThrLeuGlyLeuGluPro 357  
 DB 3421 AGCACCCAGCTCATCTACAGAGAGTGGCGTGGGAGACTACACTAGGCTGCAGCCCC 3480

OY 358 SerGluGluGluAlaProArgSerProLeuAlaProSerGlyAlaGlySerAspVal 377  
 DB 3481 TGTGAAG 3540

OY 378 PheAspGlyAspLeuGlyMetGlyAlaAlaGlyIleGlnSerLeuProThrHisAsp 397  
 DB 3541 TTTGATGGAGCTGGGGAATGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3600

OY 398 ProSerProLeuGluArgTyrSerGluAspProThrValProLeuProSerGluThrAsp 417  
 DB 3601 CCCACCCCTCTACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660

OY 418 GlyTyrValAlaProLeuThrCysSerProGluProGluTyrValAsnGlnProAspVal 437  
 DB 3661 GGCTACGTTGGCCCCCTGACCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720

OY 438 ArgProGluProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAla 457  
 DB 3721 CGGCCCCAG 3780

Oy 458 ThrleugluAAlaValThleuSerProgluYlsasnglyVallyAspValphe 477  
 Db 3781 ACTGGGAAGGCCCACTCTCTCCAGGAGAAATGGGCTGCTCAAGACCTTTT 3840  
 Oy 478 AlApheslyglYAlaValAsnProgluYrleuThrProglnglyAlaAlaPro 497  
 Db 3841 GCCTTGGGGGTCCTGGAGAACCCGAGTACTGACACCCAGGAGAGCTTC -CCT 3899  
 Oy 498 GlnProHlsProProProAlaPheSerProAlaPheAsnLeuYrTyTrpAspGln 517  
 Db 3900 CAGCCCAACCTCTCTCTGCTTACGAGCCTTCGACCAACCTTATTACTGGGACAG 3959  
 Oy 518 AsProProProgluAArglyAlaProProSerThrPheYsglyThrProThAlaGluAsn 537  
 Db 3960 GACCACACGAGCGGGGCTCCACCCAGCACTTCAAGGACACCTACGCGACAGAAC 4019  
 Oy 538 ProgluYrleuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerPro 556  
 Db 4020 CCAGAGTACTGGCTCTGGAGCTGCCA-----GTGTACACCAAGGCCAAGTCCG 4070  
 RESULT 4  
 LOCUS BM721340 653 bp mRNA linear EST 01-MAR-2002  
 DEFINITION UI-E-EO1-a1b-p-15-0-UI r1 UI-E-EO1 Homo sapiens cDNA clone  
 UI-E-EO1-a1b-p-15-0-UI 5', mRNA sequence.  
 ACCESSION BM721340  
 VERSION BM721340.1 GI:19041189  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 653)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 Tissue Procurement: Dr. Gregg Hegeman  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.  
 Location/Qualifiers  
 1. 653  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-EO1-a1b-p-15-0-UI"  
 /clone\_11b="UI-E-EO1"  
 /tissue\_type="fetal eye"  
 /dev\_stage="fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)."  
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 UI-E-EO1 is a normalized cDNA library containing the  
 following tissue(s): fetal eye. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT7T3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the

(dr)18 tail. The sequence tag for this library is  
 CGCGTATACC. This library was created for the program, Gene  
 Discovery in the Visual System, supported by National Eye  
 Institute (NEI).  
 BASE COUNT 130 a 226 c 188 g 108 t 1 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,78e-76 Length: 653  
 Score: 1087.00 Matches: 202  
 Percent Similarity: 96.19% Conservative: 0  
 Best Local Similarity: 96.19% Mismatches: 6  
 Query Match: 29.07% Indels: 2  
 DB: 14 Gaps: 1  
 US-09-821-883-2 (1-690) x BM721340 (1-653)  
 Oy 347 GlylyglYAspLeuThrleuGlyleuGluProSerGluGluAlaProArgSerPro 366  
 Db 3 GCGGGTGGGACCTGACACTAGGGCTGGAGCCCTGTGAAGAGAGGCCCCAGGCTTCCA 62  
 Oy 367 LeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAla 386  
 Db 63 CTGGCACCTCCGAGAGGGGCTGGCTCCGATATTGATGGTGACCTGGGAGGAGCA 122  
 Oy 387 AlaYsglyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgYrSerGlu 406  
 Db 123 GCAAGGGGCTGCAAGAGCTCCGCCACATGACCCACCCCTCTACACGGGTACAGTGAG 182  
 Oy 407 AsProThrValProleuProSerGluThrAspGlyYrValAlaProleuThrCysSer 426  
 Db 183 GACCCCAAGTACCCCTGCTGCTGAGACTATGGCTACGTTGGCCCCCTGACCTGCAGC 242  
 Oy 427 ProGlnProgluYrValAsnGlnProAspValArgProGlnProProSerProArgGlu 446  
 Db 243 CCCAGCCTGATATGTGAACACAGCATTTGGCCCCAGCCCTTCCGCCGAGAG 302  
 Oy 447 GlyProleuProAlaAlaArgProAlaGlyAlaThrleuGluArgAlaYrThleuSer 466  
 Db 303 GGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362  
 Oy 467 ProgluYlsasnglyVallyAspValPheAlaPheGlyAlaValGluAsnPro 486  
 Db 363 CCAGGGAAGATGGGGTGTAAAGACTTTTGGCTTTGGGGGCTGTGAAGAACCC 422  
 Oy 487 GluYrleuThrProGlnGlyAlaAlaProGlnProHisProProAlaPheSer 506  
 Db 423 GAGTACTTGACACCCAGGAGAGAGCTGCCCTCAGCCCACTCTCTGCTTACG 482  
 Oy 507 ProAlaPheAsnLeuYrTyTrpAspGlnAspProProGluArgGlyAlaProPro 526  
 Db 483 CCAGCCTTCGACCAACCTTATTACTGAGACACCCACAGAGCGGGGGCTCCACCC 542  
 Oy 527 SerThrPheYsglyThrProThAlaGluAsnProgluYrleuGlyLeuAspValPro 546  
 Db 543 AGCACTTTCANAGGAGACCTTACGCGACAGAACCCAGAGTACTGGCTGGAGCTGCCA 602  
 Oy 547 AlAAlaAlaProAlaArgSerProSerPro 556  
 Db 603 -----GTGTACACCAAGGCCAAGTCCG 626  
 RESULT 5  
 LOCUS AL701765 569 bp mRNA linear EST 22-MAR-2002  
 DEFINITION DKEZ0686P06148\_r1 686 (synonym: h1cc3) Homo sapiens cDNA clone  
 DKEZ0686P06148 5', mRNA sequence.  
 ACCESSION AL701765  
 VERSION AL701765.1 GI:19685121  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.



REFERENCE 1 (bases 1 to 569)  
 Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann  
 S.  
 TITLE EST (Duesterhoeft, et al.)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Duesterhoeft A  
 MFS

Am Klopferspitz 18a D-82152 Martinsried, Germany  
 This is the 5' sequence of the clone insert  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing  
 consortium of the German Genome Project.  
 No sl sequence available.

This clone (DKFZp686P06148) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

Location/Qualifiers  
 1..569  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp686P06148"  
 /clone\_lib="686 (synonym: h1cc3)"  
 /tissue\_type="human skeletal muscle"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Vector: pR1plex2; site\_1: SfiI; site\_2: SfiIb;  
 cDNA-collection"

BASE COUNT 115 a 187 c 156 g 110 t 1 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 2.7e-73 Length: 569  
 Score: 1046.00 Matches: 188  
 Percent Similarity: 98.95% Conservative: 0  
 Best Local Similarity: 98.95% Mismatches: 2  
 Query Match: 27.98% Indels: 1  
 DB: 9 Gaps: 0

US-09-821-883-2 (1-690) x AL701765 (1-569)

QY 127 PheGluAspAntTyrAlaLeuAlaValLeuAspAsnGlyAspProLeuAsnAsnThrThr 146  
 Db 1 TTTGAGGACAC-TATGCGCTGGCGCGGTAGACAAATGGAGACCCGCTGAACAATACACACC 59  
 QY 147 ProValThrGlyAlaSerProGlyLeuArgGluLeuGlnLeuArgSerLeuThrGlu 166  
 Db 60 CCTGTACAGGGGCGCTCCCAAGAGGCGTGGGAGGCTGACCTGGAAGCCTCACAGAG 119  
 QY 167 IleLeuLysGlyValLeuIleGlnArgAsnProGlnLeuCysTyrGlnAspThrIle 186  
 Db 120 ATCTTGAAAGAGAGGGGCTGTATCCAGCGGAACCCCGACTGCTGACACAGACAGATT 179  
 QY 187 LeuTTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsn 206  
 Db 180 TTGTGGAAAGACATCTTCACAAGAACACAGCTGGCTCTCACACTGATTAACACCAAC 239  
 QY 207 ArgSerArgAlaCysHisProCysSerProMetCysLysGlySerArgGlyTrpGlyGlu 226  
 Db 240 CGCTCTCGGGCGCTGACCCCTGTTCTCCGATGTGAAGGGCTCCCGCTGGGGAGAG 299  
 QY 227 SerSerGluAspCysGlnSerLeuThrArgThrValCysAlaGlyCysAlaArgCys 246  
 Db 300 AGTTGTGAGAGATTGTCAAGAGCTGACGCCGCACTGTCTGCGGGGTGCTGCCCGCTGC 359  
 QY 247 LysGlyProLeuProThrAspCysHisGlnGlnCysAlaAlaGlyCysThrGlyPro 266  
 Db 360 AAGGGGCGCACTGCCACTGCTGCTGAGACAGTGTGCTGCGGGCTGACGGGGCCCC 419  
 QY 267 LysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyIleGlyGluLeuHis 286  
 Db 420 AAGCACTGTGACTGCTGGCTGCTGCCCTCCACTTCAACACAGTGGCATCTGTAGCTGAC 479

QY 287 CysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGly 306  
 Db 480 TGGCCAGCCCGCTGTCACCTACACAGACAGCTTTGAGTGCATGCCCAATCCCGAGGCG 539  
 QY 307 ArgTyrThrPheGlyAlaSerCysValThr 316  
 Db 540 CGGTATACATTGCGCGCAGCTGTGACT 569

## RESULT 6

BO958632

LOCUS BO958632 943 bp mRNA linear EST 21-AUG-2002  
 DEFINITION AGNC000RT\_10061234 NCI\_CGAP\_Mam2 Mus musculus cDNA clone  
 IMAGE:6485621 5', mRNA sequence.

ACCESSION BO958632  
 VERSION BO958632.1 GI:22374110

KEYWORDS EST.  
 SOURCE house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE NIH Unpublished (1999)  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs.femail.nih.gov

Tissue procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.llnl.gov

Plate: LLM14027 row: 1 column: 06  
 High quality sequence stop: 579.

## FEATURES

source

Location/Qualifiers  
 1..943  
 /organism="Mus musculus"  
 /strain="FVB/N-3"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6485621"  
 /clone\_lib="NCI\_CGAP\_Mam2"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-Sport6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"

## BASE COUNT

203 a 312 c 249 g 179 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.85e-67 Length: 943  
 Score: 976.00 Matches: 187  
 Percent Similarity: 77.04% Conservative: 11  
 Best Local Similarity: 72.76% Mismatches: 43  
 Query Match: 26.10% Indels: 16  
 DB: 14 Gaps: 2

US-09-821-883-2 (1-690) x BO958632 (1-943)

QY 322 TyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisAspHisArgSer 341  
 Db 18 TTCTTCTCCCAAGACCTGCGCTAGTACTGGAGACACAGCCACGACAGACCGCAGC 77  
 QY 342 SerSerThrArgSerGlyGlyLysAspLeuThrLeuGlyLeuGluProSerGluGluGlu 361  
 Db 78 TCGTGGCGAGAGAGGGGGTGTGAGCTGACACTGGGGCTGGAGCCCTCCGAAAGAG 137  
 QY 362 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspLysp 381  
 Db 138 CCCCCAGATCTCCACTGCTGCCCTCCCGAAGGGGCTGCTCGATGTGTGATGTGTGAC 197



```

VERSION      AM370693.1  GI:6875347
KEYWORDS
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 614)
AUTHORS      Mammalia; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
TITLE        Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
JOURNAL      HCGP http://www.ludwig.org.br/ORESTES.
COMMENT      The FAPESP/LICR Human Cancer Genome Project
              Unpublished (1999)
              Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL.
              (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1st2-QV1-BT0260-
              011199-024-all&t3=1999-11-01&t4=1)
              Seq primer: puc 18 forward
              High quality sequence start: 12
              High quality sequence stop: 612.
              Location/Qualifiers
                1..614
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone_id="BT0260"
                  /dev_stage="Adult"
                  /note="Organ: breast; Vector: puc18; Site.1: Sma1; Site.2:
                  Sma1; A mini-library was made by cloning products derived
                  from ORESTES PCR (U.S. Letters Patent application No. 196
                  ,716 - Ludwig Institute for Cancer Research) profiles
                  into the puc 18 vector. Reverse transcription of tissue
                  mRNA and cDNA amplification were performed under low
                  stringency conditions."
BASE COUNT   119 a 198 c 172 g 124 t 1 others
ORIGIN
Alignment Scores:
Pred. No.:      7.5e-63      Length:      614
Score:          915.00      Matches:      158
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    24.47%      Indels:      0
                  Gaps:      0
US-09-821-883-2 (1-690) x AM370693 (1-614)
Qy 172 ValLeuIleGlnAArgAsnProGlnLeuGlyGlnAspThrIleLeuTrpLysAspIle 191
Db 17 GTCTTGATCCAGCGGACCCAGCTCTGCTACACGAGACGATTTTGTGAAAGACATC 76
Qy 192 PheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgLacys 211
Db 77 TTCCCAAGAACACACACCTGCTCTCACACTGATAGACACCAACCCGCTCGGGCTTGC 136
Qy 212 HisProCysSerProMetCysLysGlySerArgCysTrpGlyGlnLysSerGlnLysPcys 231
Db 137 CACCCCTGTCTCCGATGTGTAAAGGGCTCCGCTGCGGAGAGATTCTGAGAGATTGT 196
Qy 232 GlnSerLeuThrArgThrValCysAlaLysGlyCysAlaAlaArgCysLysGlyProLeuPro 251
Db 197 CAGACCCCTGAGCGGACGCTGTGTGCGGTGCTGTGCGGCTGCAAGGGGCGACTTGGCC 256
Qy 252 ThrAspCysCysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCys 271
Db 257 ACTGACTCTGCTGCATGACAGAGTGTCTGCGGCTGCACGGGCGCCAAACACTCTGACTGC 316
Qy 272 LeuAlaCysLeuHisPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuVal 291

```

```

Db 317 CTGGCGTCTCTCCACTTCACACAGATGCGATCTGTGAGCTGCACATGCCAGCCCTGGTC 376
Qy 292 ThrTyraThrAspThrPheGlnLysSerMetProAsnProGlnLysGlyArgTyrThrPheGly 311
Db 377 ACCATCAACACAGACACCTTTGAGTCTCCATGCCATCCCGGCGGCGGTATATCTGCGC 436
Qy 312 AlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySer 329
Db 437 GCCAGCTGTGTACTGCTGCTGTCCTACACTACTCTTTCTACGGAGCTGGATCC 490
RESULT 9
B1154872 791 bp mRNA linear EST 05-JUL-2001
LOCUS 602902857F1 NCL_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032238 5',
DEFINITION mRNA sequence.
ACCESSION B1154872 GI:14614873
VERSION B1154872
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 791)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Lotzhar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.llnl.gov
          Plate: LLM11090 row: b column: 15
          High quality sequence stop: 764.
          Location/Qualifiers
            1..791
              /organism="Mus musculus"
              /strain="129, C57BL/6J, FVB/N"
              /db_xref="taxon:10090"
              /clone_id="IMAGE:5032238"
              /clone_11b="NCI_CGAP_Mam3"
              /tissue_type="tumor, gross tissue"
              /dev_stage="10 months"
              /lab_host="DH10B"
              /note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
              Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
              Library constructed by Life Technologies. Investigators
              providing samples: Lotzhar Hennighausen/Chu-Xia Deng, NIH
              Reference for transgenic model: Xu et al., Nature Genetics
              22, 37-43 (1999)."
BASE COUNT 159 a 264 c 215 g 153 t
ORIGIN
Alignment Scores:
Pred. No.:      1.79e-62      Length:      791
Score:          912.00      Matches:      176
Percent Similarity: 82.30%  Conservative: 10
Best Local Similarity: 77.88%  Mismatches: 39
Query Match:    24.39%      Indels:      2
                  Gaps:      0
US-09-821-883-2 (1-690) x B1154872 (1-791)
Qy 322 TyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHisArgHisArgSer 343
Db 111 TTCTTCTCCGACGACCCGCTAGGTCTGAGGACAGACGACGACGACGACGACGACG 170
Qy 342 SerSerThrArgSerGlyGlyAspLeuThrLeuGlnGlyLeuGlnProSerGlnGln 361
Db 171 TCGTCGCGCCAGGAGTGTGCTGAGCTGACACTGCGGCGCTGAGAGCCCTCGAAGAGAG 230

```





```

/clone="NT2RM2001211"
/clone_lib="NT2RM2"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/notes="Vector: PMEL18FL3; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT      137 a      224 c      198 g      123 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:      4,31e-60      Length:      685
Score:          881.00      Matches:      170
Percent Similarity: 91.89%      Conservative: 0
Best Local Similarity: 91.89%      Mismatches: 8
Query Match:      23,56%      Indels:      7
DB:              9      Gaps:      1

US-09-821-883-2 (1-690) x AW123871 (1-685)
QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHis 337
DB 137 TGTCCA-----GACCCCTGCCCCGGGGCGCTGGGGCATGTCCACCAC 178
QY 338 ArgHisArgSerSerThrArgSerGlyGlyValAspLeuThrLeuGlyLeuGluPro 357
DB 179 AGGCACCGCAGCTCATCTACAGAGAGTGGCGGGAGACCTGACACTAGGGCTGGAGCCC 238
QY 358 SerGluGluGluAlaProArgSerProLeuAlaProSerGlyAlaGlySerAspVal 377
DB 239 TCTGMAAGAGAGGGCCCCAGAGTCTCCACTGCACTCCGAAAGGGGCTGGCTCGATCTA 298
QY 378 PheAspGlyAspLeuGlyMetGlyAlaAlaValGlyLeuGlnSerLeuProThrHisAsp 397
DB 299 TTTCATGATGATCACTGGGAATAGGGGGGAGCCCAAGGGGCTGCAAAAGCTCCACACATGAC 358
QY 398 ProSerProLeuGlnArgTyrSerGlyAspProThrValProLeuProSerGluThrAsp 417
DB 359 CCCAGCCCTTACAGCCGTCATGAGGAGGCCACACATACCTCCCTCTGAGACTGAT 418
QY 418 GlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAlaGlnProAspVal 437
DB 419 GGCAGCTGTCGCCCTGACCTGACGCCGCCAGCTCAATATGTGAACCAAGCCAGATGTT 478
QY 438 ArgProGlnProProSerProArgGlyGlyProLeuProAlaAlaArgProAlaGlyAla 457
DB 479 CGGCCCCAGCCCTTGGCCCCGAGAGGGCTCTGCTCTGCCCCAGCTGCTGGTGC 538
QY 458 ThrLeuGluArgAlaTyrThrLeuSerProGlyLysAsnGlyValValLysAspValPhe 477
DB 539 ACTCTGMAAGAGGCCAAGACTCTCTCCCAAGGAGAAATGGGGTGGTCAAAAGACGTTTTT 598
QY 478 AlaPheGlyGlyAlaValAlaGluAsnProGluTyr-LeuThrProGlnGlyGlyAlaAlaPro 497
DB 599 GCGTTTGGGGTGGCTGGGAGAAACCCGAGTACTTGTACACCCCAAGAGAGAACTGCCCC 658
QY 497 OGlnProHisPro 501
DB 659 CTTAAAGCCACACC 671

RESULT 13
AW701942      609 bp      mRNA      linear      EST 18-APR-2000
LOCUS      uq93802.y1 NCI_CGAP_Mam10 Mus musculus cDNA clone IMAGE:2938635 5'
DEFINITION      similar to gb:M1130 ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE
PRECURSOR (HUMAN); mRNA sequence.
ACCESSION      AW701942
VERSION      AW701942.1 GI:7586086
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 609)

```

```

AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL      Tumor Gene Index
COMMENT      Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Jeffrey Green M.D.,
Gilbert Smith, Ph.D., William Muller, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.lnl.gov/image/html/iresources.shtml
MGI:1051055
Seq primer: -40RP from GIBCO
High quality sequence stop: 512.
FEATURES
source
1. 609
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:2938635"
/clone_lib="NCI_CGAP_Mam10"
/tissue_type="pooled mammary gland tumors"
/lab_host="Dh10B (phage-resistant)"
/notes="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia
) with a modified polylinker; Site.1: NotI; Site.2: EcoRI;
1st strand cDNA was prepared from mRNA obtained from
pooled mammary gland tumors with a Not I - oligo(dT)
primer (5'
TGTACCAATCTGAAGTGGAGGCGGCACATGTTTTTTTTTTTTTTT 3').
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      125 a      195 c      159 g      127 t      3 others
ORIGIN

Alignment Scores:
Pred. No.:      4e-59      Length:      609
Score:          866.00      Matches:      163
Percent Similarity: 85.07%      Conservative: 8
Best Local Similarity: 81.09%      Mismatches: 30
Query Match:      23,21%      Indels:      0
DB:              10      Gaps:      0

US-09-821-883-2 (1-690) x AW701942 (1-609)
QY 346 SerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluAlaProArgSer 365
DB 1 AGTGGAGGTGTGAGTGCACATGGGCTGGAGCCCTCGGAAGAGGGCCCCCAGATCT 60
QY 366 ProLeuAlaProSerGlyGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGly 385
DB 61 CCACGTGCTCCCTCGGAAGGGGCTGGCTCGATGTGTTATGTGTGACCTGCCAAATGGGG 120
QY 386 AlaAlaGlyGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSer 405
DB 121 GTAAACCAAGGGCTGCAGAGCTCTCTCCACATGACTCAGCCCTCTACAGCGGTACAGC 180
QY 406 GluAspProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCys 425
DB 181 GAGGAGCCCACTTACTCTGCCCCCGAGACAGATGAGGTGTGCTGCCCGGGCTGC 240
QY 426 SerProGlnProGluTyrValAlaGlnProAspValArgProGlnProProSerProArg 445
DB 241 AGCCCCAGCCCGAGATGTGAACCAATCAAGAGTTCAGCTCAGGCTCTTAAACCCA 300
QY 446 GluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeu 465

```

Db	301	GAGGTCCTCTGCGCCCTC	CGTCGCGGCTCTGCTACTCTAGAAAGCAAGCAAGCTCTC	360
Qy	466	SerProGlyLyAsnGlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsn		485
Db	361	TCCTCTGGAGAAAGGGGCTGTCTAAACACGTTTTTGGCTTCGGGGGCTGTGTGGAGAAC		420
Qy	486	ProGluTyrLeuThrProGlnGlyValAlaProGlnProHisProProProAlaPhe		505
Db	421	CCTGATACTTAGTACCCAGAGAAAGGCACTGCTCTCCGCCCCACCTCTTCCTCTCCCTTC		480
Qy	506	SerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProProGluArgGlyAlaPro		525
Db	481	AGCCACGCTTTGACACACCTCTATTACGTGNGACCAAGACTATCGAGACAGAGNGCTCCA		540
Qy	526	ProSerThrPheLeuGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspVal		545
Db	541	CCAACTACTTTGAAGGACCCCTCCCTCAGAGAACCTGAGTACTTGGCTGGATGTA		600
Qy	546	Pro	546	
Db	601	CCT	603	
RESULT 14				
LOCUS	BC283493	621 bp	mRNA	linear
DEFINITION	602407782p1 NIH_MGC_91 Homo sapiens CDNA clone IMAGE:4519652 5',			
ACCESSION	BC283493			
VERSION	BC283493.1	GI:13033480		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS	1 (bases 1 to 621)			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999)			
	Contact: Robert Strusberg, Ph.D.			
	Email: cga@ds-remail.nih.gov			
	Tissue Procurement: DCTD/DBP			
	cDNA Library Preparation: Life Technologies, Inc.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:			
	http://image.lnl.gov			
	Plate: L14M10416	row: d	column: 21	
	High quality sequence stop.			
FEATURES				
source	1..621			
	location/Qualifiers			
	1..621			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:4519652"			
	/clone_lib="NIH_MGC_91"			
	/tissue_type="adenocarcinoma, cell line"			
	/lab_host="DH10B (phage-resistant)"			
	/note="Organ: prostate; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."			
BASE COUNT	131 a	207 c	178 g	104 t
ORIGIN				1 others
Alignment Scores:				
Pred. No.:	7.78e-59	Length:	621	
Score:	864.50	Matches:	174	
Percent Similarity:	91.67%	Conservative:	2	
Best Local Similarity:	90.62%	Mismatches:	12	
DB:	23.12%	Indels:	7	
	12	Gaps:	1	

OY	371	GIUGIYALAGIYSERASPVAlpheaspGLyspleuglmetGtlyAlaAlayGLyLeu	390
Db	2	GAAAGGGCTGCCTCCGATGTATTGTATGCTACCTGGGAAGGGGGCGCAAGGGGCTG	61
OY	391	GInserLeuProThrHISAsPProSerProLeuGlnAgtTyrSerGluAsPProThrVal	410
Db	62	CMAAGCCTCCCACATGACGCCAGCCCTTACAGCGGTACTGAGAGACCACACATA	121
OY	411	ProLeuProSerGluTrHisPGLYTYrValAlaProLeuThrCySserProGlnProGlu	430
Db	122	CCCCGCCCTGTGACATGATGGCTACCTTGCCCCCTGACCTCACGCCCCACACTGAA	181
OY	431	TyrValaIngInProAsPVAlargProGlnProProSerProArGluGluProLeuPro	450
Db	182	TATGTGAACCAAGCCAGATGTTGGCCCAAGCCCTTGCCCCGAGAAGGCCCTTGCT	241
OY	451	AlaAlaArProAlaGlaAlaThrLeuGluLuarGalaLystrHLeuSerProGlyLysAn	470
Db	242	GCTGCCAGACCTGCTGTGGCCACTCTGGAAAGGCCCAAGACTCTCTCCCAAGGAAGAT	301
OY	471	GLYVALIYLVSASPVALPhealaphedLYglYAAlaValIGluANSPROGLTYrLeuThr	490
Db	302	GGGGTCGTCAAAGAGTTTTTGCTTTGGGGGTGGCGTGGAGAACCCGAGTACTTGACA	361
OY	491	PROGLINGLYGYAlaAlaProGln-ProHISProProProAlaPheserProAlaPhas	510
Db	362	CCCCA-GGAGAGAGCTNGCCCTCAGCCCAACCTCTCTCTGCTTACACCCAGCTTGA	420
OY	510	PasnLeuTyrrTYrTPasPGLInAsPProProGluArgLYAlaProProSerThrPhely	530
Db	421	CAACCTCTATTACTGGACCAAGGA-CCACACAGAGGGGGGCTCCACCCACACAC-TTCA	478
OY	530	sGLYrPrOmThraGluaANSPROGLTYrLeuGlyLeuAsPVAlproAla-----	547
Db	479	AGGGACACCTTAGGGAGAGAACCCAGAGTACTGGGTGTGGACGTGCCAGTGTGAACAG	538
OY	548	-AlaAlaProAlaArGerProSerProSerThr	558
Db	539	AAGCAAGTCGCCAGAACCCCTGATGTCTCTCA	572
RESULT 15			
BOJ17097			
LOCUS	BOJ17097	887 bp	mRNA linear EST 16-JUL-2002
DEFINITION	AGENCOURT_8100479 Lupsk1_symplethetic_trunk Homo sapiens cDNA clone		
IMAGE:	6190537 5', mRNA sequence.		
ACCESSION	BOJ17097		
VERSION	BOJ17097.1	GI:2185594	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs@email.nih.gov Tissue Procurement: Dr. James R. Lupski cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LMN) Clone distribution: MGC Clone distribution information can be found through the I.M.A.G.E. Consortium/LMN at: http://image.llnl.gov Plate: LMN13589 row: 1 column: 02 High quality sequence stop: 556. Location/Qualifiers 1..887 /Organism="Homo sapiens" /db_xref="taxon:9606"		





GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 7, 2003, 22:19:33 ; Search time 1592 seconds

(Without alignments)  
976.055 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAPLLARASLSLGLF.....EPVQEGAPPPRAAHNNHHH 690

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-rlp  
-Q-/cgn2\_1/USPFO.spool/US09821883/runat\_28042003\_104748\_14558/app\_query.fasta.1.839  
-DB-N Geneseq.101002 -QFMT-fastap -SUFFIX-p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE-PCT -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT-ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09821883 @CGN.1.1.396 @runat\_28042003\_104748\_14558 -NCPU=6 -ICPU=3  
-NO\_XLPRY -NO\_MMAR -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-WARN\_TIMEOUT=30 -THRAVS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Geneseq.101002:\*

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*

2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*

3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*

4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*

5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*

6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*

7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*

8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*

9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*

10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*

11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*

12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*

13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*

14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*

15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*

16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*

17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*

18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*

19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*

20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*

21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*

22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*

23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*

24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3739	100.0	2070	22 AAD21565	Human HER500-hgm-C
2	3473.5	92.9	2091	22 AAD21567	Human HER500-rcm-C
3	2974	79.5	1665	22 AAD21564	Human HER500 fusio
4	2959.5	79.2	1692	22 AAD21566	Human HER500 fusio
5	2423	64.8	4530	16 AAT01585	Her-2/neu (ERBB2)/c
6	2423	64.8	4530	18 AAT71253	Human HER2 gene.
7	2423	64.8	4530	21 AA260815	Nucleotide sequenc
8	2423	64.8	4530	22 AAD19731	Human HER2-tyrosine kin
9	2423	64.8	4530	22 AAD19731	Human HER2-tyrosine kin
10	2423	64.8	4530	24 ABR83918	Human HER2-tyrosine kin
11	2422	64.8	4530	24 ABR83918	Human HER2-tyrosine kin
12	2418	64.7	4472	21 AA14812	Human HER2 (ETB2)
13	2418	64.7	4473	20 AA231071	CDNA encoding the
14	2418	64.7	4473	24 AAD38904	HER-2 nucleic acid
15	2415	64.6	4274	22 AAF24297	Human Her-2 DNA.
16	2410.5	64.5	3768	21 AA09455	HER2 transgene pla
17	2410.5	64.5	3768	21 ABL91709	Human heregulin 2
18	2410.5	64.5	3768	24 ABL91709	Human polynucleoti
19	2405.5	64.3	3768	17 AAT40739	Human HER2 (ETB2)
20	2405.5	64.3	3768	20 AA01912	HER-2/neu oncogene
21	2405.5	64.3	3768	22 AAD23392	Human HER-2/neu on
22	2405.5	64.3	3768	24 AAD22743	Human HER-2/neu pr
23	2405.5	64.3	3768	24 ABA92250	Human Her-2/neu pr
24	2405.5	64.3	3768	24 ABL10730	Human Her-2/neu DN
25	2336.5	64.1	4299	14 AA046083	Sequence encoding
26	2334.5	62.4	3678	24 ABR66207	CDNA encoding huma
27	2258.5	60.4	2385	18 AAT72725	Her2-GM-CSF immuno
28	2188.5	58.5	1437	22 AAD21568	Human HER300-rcm-C
29	2130	57.0	2781	24 ABA92253	Human HER-2/neu ex
30	2129	56.9	2763	24 ABA92252	Human HER-2/neu ex
31	2092.5	56.0	3600	21 AA089736	Human HER-2/neu co
32	1968	52.6	3955	16 AAT01590	Rat HER-2/neu prom
33	1968	52.6	3955	21 AA089736	Rat HER-2/neu prom
34	1961	52.4	3771	21 ABA82210	Mouse Her-2/neu cd
35	1961	52.4	3771	22 ABA82210	Mouse Her-2/neu cd
36	1961	52.4	3771	22 ABA82210	Mouse Her-2/neu cd
37	1680.5	44.9	2871	21 AA250586	Nucleotide sequenc
38	1614.5	43.2	1872	11 AA006828	DC83cFv-erbB2EC fu
39	1587	42.4	867	22 AAD21571	Extracellular port
40	1199.5	32.1	1191	22 AAD21573	Mature human HER-2
41	1188	31.8	1806	24 AAD32745	Human cDNA for the
42	1183	31.6	1755	24 AAD32746	Human cDNA for the
43	1183	31.6	1767	24 AAD32744	Human cDNA for the
44	1183	31.6	1773	24 AAD32747	Human cDNA for the
45	1182	31.6	651	22 AAD21572	Mature human HER-2

## ALIGNMENTS

RESULT 1

AAD21565

28-JAN-2002 (first entry)

Human HER500-hgm-CSF fusion DNA construct.

Immunostimulatory fusion protein: IRP; antigen component; therapy;

Immunostimulatory component; T-cell mediated immune response; DC;

dendritic cell; colon cancer; breast carcinoma; ovarian cancer;

PAP protein: Ala Arg linker; membrane distal extracellular domain;

membrane distal intracellular domain; C-terminal tag; human; GM-CSF;

HER-2 protein; granulocyte-macrophage colony stimulating factor;

HER500-hgm-CSF fusion DNA; ds.

```

OS Chimeric - Homo sapiens.
OS Chimeric - Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..2070
FT     /product= "Human HER500-hgm-CSF fusion protein construct
FT     /tag= a comprising human PAP signal sequence, mature PAP protein
FT     /note= "an Ala Arg linker, human HER-2 signal sequence,
FT     mature HER-2 membrane distal extracellular and
FT     intracellular domains, an Ala Ala linker, a mature
FT     human GM-CSF sequence and a C-terminal tag"
FT     /note= "CDS does not include stop codon"
FT     /partial
XX
PN WO200174855-A2.
PD
PP 11-OCT-2001.
XX
PE 30-MAR-2001; 2001WO-US10515.
XX
PR 30-MAR-2000; 2000US-193504P.
XX
PA (DEND-) DENDREON CORP.
XX
PI Laus R, Vidovic D, Graddis T;
XX
DR WPI: 2001-662965/76.
XX
DR P-PDB: AAEI3109.
XX
PT An immunostimulatory fusion protein comprising the intracellular domain
PT of HER-2 and an antigen elicits an immune response to the antigen and
PT is useful for the treatment of associated cancer associated -
XX
PS Example 3; Page 28; 59pp; English.
XX
CC The invention relates to immunostimulatory fusion proteins (IFP) and
CC nucleic acid molecules encoding such proteins. The IFPs comprise a
CC polypeptide antigen component and an immunostimulatory component derived
CC from the intracellular domain of HER-2 protein which is effective to
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFP or superactivated dendritic cells
CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC hgm-CSF fusion DNA construct which comprises DNA molecules encoding
CC human PAP signal sequence, mature PAP protein, an Ala Arg linker, human
CC HER-2 signal sequence, mature HER-2 membrane distal extracellular and
CC intracellular domains, an Ala Ala linker, a mature human granulocyte-
CC macrophage colony stimulating factor (GM-CSF) sequence and a
CC C-terminal tag.
XX
XX Sequence 2070 BP; 425 A; 707 C; 568 G; 370 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 7, 02e-156 Length: 2070
XX Score: 3739.00 Matches: 690
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-821-883-2 (1-690) x AAD21565 (1-2070)
XX
OY 1 McArGaAlAaIProLeuLeuLeuAlAARGAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
Db | |||||
OY 1 ATGAGACGTGCACCCCTTCCTCTCGTGGCAAGGACCAAGCCTTGACCTTGCTTTGTTT 60
Db | |||||
OY 21 LeuLeuPhePheTPLeuAsPARGSerValLeuAlaLysGLuLeuAlaArgGLyAlaAla 40
Db | |||||
OY 61 CTGCTTTTTTTCTGGCTAGACCAAGGTGTACTAGCCAAGGATTTGGCGCGGGGCCGCG 120
Db | |||||
OY 41 SerThrGlnValLCysTrnGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db | |||||

```

D	121	TCGACCCAGTGTCCACCGGCACAGACATGAACTGGGCTCCCTCCAGCTCCGACAGC	180
Q	61	HisLeuaspMetLeuAtrgHisLeuYrgInglyCysGlnValValIcInLysAnLeuGlu	80
D	181	CACCTGGACATATGCTCCGCCACCTTCACAGAGGCTGCCAGGTGGTGCAGGAACCTGGAA	240
Q	81	LeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIleValGln	100
D	241	CTCACTTACCTGGCCACCAATGGCAGCGCTCTCTCTCTCGAGATATCCAGAGGGTGCAG	300
Q	101	GlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnIrrgLeuArgIle	120
D	301	GGCTACGGCGCATCGCTTCACAAACAAAGTAGGACAGTCCCACTGAGAGCGCTGGGAAATT	360
Q	121	ValArgGlyThrGlnLeuPheGlnAspAsnTrpAlaLeuAlaValLeuAspAsnLysAsp	140
D	361	GTGGAGGACCCACCGCTTTGTGAGGACAACTTAGCCCTGGCGGTGCTATACATAGGAGAC	420
Q	141	ProLeuAsnAsnTrpThrProValIrrhGlyAlaSerProGlyGlyLeuArgIleLeuGln	160
D	421	CCGCTGAACATATTCACCCTGTGCACAGGGGGCTCCCGCAGAGGCGCTGGGGAGCTGCAG	480
Q	161	LeuArgSerLeuThrGlnIleLeuLysGlyGlyValIleuIleGlnArgAsnProGlnLeu	180
D	481	CTTGGAAGCCTCACAGAGACTTTGAAAGAGGGGCTTTATTCAGCGGAACCCGCAAGCTC	540
Q	181	CysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnIleuAlaLeu	200
D	541	TGCTACACAGACACAGATTTGTGGAGGACATTTCCACAAACACACAGCTGGCTCTTC	600
Q	201	ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly	220
D	601	ACACTGATAGACACAAACCCCTCTCGGGCTGCACCCCTGGTCTCCGATGTGTAAGGAC	660
Q	221	SerArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgTrpValCysAla	240
D	661	TCCCGCTGCTGGGAGAGAACTTCTGAGAAATGTTCACAGCTCAGCGCACTCTCTGTGCC	720
Q	241	GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnIleCysAla	260
D	721	GGTGGCTGTGGCCGCTGCAGAGGGGCCATGCCACATGACTGGTCCCATAGCAGTGTCT	780
Q	261	AlaGlyArgThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer	280
D	781	GCCGGCTGCACGGGCCCAAGCACTGACTGCTGGCTGGCTCCACTCAACACACAGT	840
Q	281	GlyIleCysGlnLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGlnSer	300
D	841	GGCATCTGTAGCTGCACCTCCCAAGCCCTGTCTACTTACACACAGACACAGCTTTAGTCC	900
Q	301	MetProAsnProGlnGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTrp	320
D	901	ATGCCCAATCCCGAGGGCCGGTATATCATTTCCGGCGCAGCTGTGTACTCTCTCTTAC	960
Q	321	AsnTrpLeuSerThrAspValGlySerGlyAlaGlyIleValHisAsnArgHisArg	340
D	961	AACATACCTTTCTACGGACGTGGGATCGGGCGCTGGGGCATGTGTCCACACAGGCACGC	1020
Q	341	SerSerSerThrArgSerGlyGlyLysAspLeuThrLeuGlyLeuGlnProSerGlnGlu	360
D	1021	AGCTCATCTTCCACAGAGATGGCGGTGGGACCTGTACACTAGGGCTGGAGCCCTCTTAAG	1080
Q	361	GlnAlaProArgSerProLeuAlaProSerGlnGlyAlaGlySerAspValPheAspLys	380
D	1081	GAGGCCCCACAGTCTCCACATCGACACCTCCGAAGGGGCTGGCTCCGATGTATTGATGTT	1140
Q	381	AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro	400
D	1141	GACCTGGGAATGGGGGACGACCAAGGGGCTGCAAAGGCTCCACACATATGACCCAGCCT	1200
Q	401	LeuGlnArgTrpSerGlnAspProThrValProLeuProSerGlyThrAspLysIleVal	420
D	1201	CTTACAGGGGTACAGTAGAGCCCCACAGTACACCTTCCCTTCTGAAAGATGATGGTCTACGTT	1260

OY	421	AlAProLeuThrCySerProGlnProGluTyValAsnGlnProAspValArgProGln	440
Db	1261	GGCCCCCTGACCTGCACCCCGACCGCTGAATGTAACACGACGAGATGTTGGCCGAG	1320
OY	441	ProProSerProAArgGluGlyProLeuProAlaAlaArgProAlaGlyValaThrLeuGlu	460
Db	1321	CCCCCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCTGCCGACCTGTGTGCTCGGAA	1380
OY	461	ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly	480
Db	1381	AGGGCCAAAGACTCTCTCCCGAGGAGAAATGGGGTGCCTCAAAAGACTTWTTCCTTGGG	1440
OY	481	GlyAlaValGluAsnProGluTyLeuThrProGlnGlyValAlaLeuProGlnHis	500
Db	1441	GGTCCCGGAGAAACCCCGAGTACTGTACACCCCGAGGAGGAGCTGCCCTCAGGCCAC	1500
OY	501	ProProProAlaPheSerProAlaPheAspAsnLeuTyThrTyTrpAspGlnAspPro	520
Db	1501	CCCTCCTCCTGCTTCACGCCAGCCTTGAGAACCTCTATTACGGGACAGGACCCACCA	1560
OY	521	GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTy	540
Db	1561	GAGCGGGGGGCTCCACCCACACCTTCMAAGGAGACCTACGGCAGAGAACCCAAAGTAC	1620
OY	541	LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSerThrGlnPro	560
Db	1621	CTGGGTCTGGACGCGCCAGCGGCGGACCCGCGCTCCGCCACGCCACAGACGCC	1680
OY	561	TrpGlnHisValAsnAlaAlaGlnGlyAlaArgArgLeuLeuAsnLeuSerArgAspThr	580
Db	1681	TGGGAGCATGTGATGGCATTCAGAGAGCCCGGCGCTCTCTGAACTGATGAGACACT	1740
OY	581	AlaAlaGluMetAsnGluThrValGluValIleSerGluMetPheAspLeuGlnIlePro	600
Db	1741	GCTCTGAGATGATGAACAGATGAAGTATCTCAGAAATGTTTACCTCAGAGAGCCG	1800
OY	601	ThrGlyLeuGlnThrArgGluGluLeuTyLysGlnGlyLeuArgGlySerLeuThrIlys	620
Db	1801	ACCTGCTACAGACCCCTGAGCTGTATACACAGGCGCTGGGGGAGGCTCCACCAAG	1860
OY	621	LeuLysGlyProLeuThrMetMetAlaSerHisTyLysGlnHisLysProProThrPro	640
Db	1861	CTCAAGGGCCCCCTTGACCATGATGGCCAGCCACTACAAACAGACATGCCCTCAACCCCG	1920
OY	641	GluThrSerCyAlaIleArgIleIleIleThrPheGluSerPheLysGluAsnLeuLysAsp	660
Db	1921	GAACCTCTCTGTGCACCCAGATTATCACCTTTGAAAGTTTCAAAAGAACCTGGAAGAC	1980
OY	661	PheLeuLeuValIleProPheAspCyTrpGluProValGlnGluGlyAlaProProPro	680
Db	1981	TTTCTGCTGTGCATCCCTTTGACTGTGGAGACCGACGTCCAGAGAGGCGCCGACCCCG	2040
OY	681	ProAlaAlaAlaHisHisHisHisHisHis 690	
Db	2041	CCGGCGGCGGCACATCACATCACATCAC 2070	
RESULT 2			
AAAD21567	AAAD21567 standard; DNA; 2091 BP.		
XX	AAAD21567:		
XX	28-JAN-2002	(first entry)	
XX	Human HER500-rGM-CSF fusion DNA construct comprising OVA-derived peptide		
KW	Immunostimulatory fusion protein: IFP, antigen component; therapy:		
KW	immunostimulatory component: T-cell mediated immune response; DC;		
KW	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;		
KW	PAP protein; Ala Arg linker; membrane distal extracellular domain;		
KW	membrane distal intracellular domain; C-terminal tag; human; GM-CSF;		
KW	HER-2 protein; granulocyte-macrophage colony stimulating factor;		

XX	ovalbumin-derived octapeptide; OVA; rat; HER500-rgm-CSF fusion DNA; ds.
XX	Chimeric - Homo sapiens.
OS	Chimeric - Rattus norvegicus.
OS	Chimeric - Unidentified.
XX	
FT	Key
FT	CDS
FT	1. 2091
FT	location/Qualifiers
FT	/*tag= a
FT	/product= "Human HER500-rgm-CSF fusion protein construct
FT	comprising human PAP signal sequence, mature PAP protein,
FT	an Ala Arg linker, human HER-2 signal sequence,
FT	mature HER-2 membrane distal extracellular domain,
FT	an Ala linker, an ovalbumin (OVA)-derived
FT	immunodominant octapeptide, HER-2 membrane distal
FT	intracellular domain, an Ala Ala linker, a mature rat
FT	GM-CSF sequence and a C-terminal tag"
FT	/note= "CDS does not include stop codon"
FT	/partial
XX	
XX	MO200174655-A2.
PN	
PD	11-OCT-2001.
XX	
XX	30-MAR-2001; 2001WO-US10515.
PF	
XX	30-MAR-2000; 2000US-193504P.
PR	
XX	(DEN2-) DENBREON CORP.
PA	
XX	Laus R, Vidovic D, Graddis T;
XX	
PI	
XX	WPI: 2001-662965/76.
DR	P-PSDB: AAEI3111.
XX	
PT	An immunostimulatory fusion protein comprising the intracellular domain
PT	of HER-2 and an antigen elicits an immune response to the antigen and
PT	is useful for the treatment of associated cancer associated -
XX	
PS	Example 1; Page 29; 59pp; English.
XX	
CC	The invention relates to immunostimulatory fusion proteins (IFP) and
CC	nucleic acid molecules encoding such proteins. The IFPs comprise a
CC	polypeptide antigen component and an immunostimulatory component derived
CC	from the intracellular domain of HER-2 protein which is effective to
CC	elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC	immune response to the antigen. IFP or superactivated dendritic cells
CC	are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC	associated with a particularly antigen. The present sequence is HER500
CC	rgm-CSF fusion DNA construct which comprises DNA molecules encoding
CC	human PAP signal sequence, mature PAP protein, an Ala Arg linker, human
CC	HER-2 signal sequence, mature HER-2 membrane distal extracellular domain,
CC	an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
CC	HER-2 membrane distal intracellular domain, an Ala Ala linker, a mature
CC	rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence
CC	and a C-terminal tag.
XX	
XX	Sequence 2091 BP; 447 A; 699 C; 559 G; 386 T; 0 other;..
SO	
XX	
XX	Alignment Scores:
XX	Prod. No.: 3,17e-144
XX	Score: 3473.50
XX	Length: 2091
XX	Matches: 641
XX	Percent Similarity: 94.42%
XX	Conservative: 19
XX	Best Local Similarity: 91.70%
XX	Mismatches: 28
XX	Query Match: 92.90%
XX	Indels: 11
XX	Gaps: 2
DB:	
DB:	US-09-821-883-2 (1-690) x AAD21567 (1-2991)
OY	1 MetArgAlaIaPProLeuLeuLeuAlaArgAlaIaIaSerLeuSerLeuGlyPheLeuPhe 20
DB	1 ATGAGAGGTGCACCCCTTCCTGCGCAGGCGCAGCAAGCCTTAGCCCTTCTTGT 60

Oy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaIysGluLeuAlaArgGlyAlaAla 40  
 Db 61 CTGCTTTTTCCTGGCTGACCGAAGTGTACTAGCCAAAGAGTTGGCGGCGCGCGG 120  
 Oy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60  
 Db 121 TCGACCCAAAGTGTGACGGGCGACAGACATGAAGTGGCGCTCCCTGCCAGTCCGAGACC 180  
 Oy 61 HisLeuAspMetLeuArgHisLeuTrpGlnIysGlnValValGlnIysAsnLeuGlu 80  
 Db 181 CACCTGGACATGCTCCGCCACCTCTACACAGGGCTCCAGAGTGGTCCAGGAAACCTGGAA 240  
 Oy 81 LeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIValGln 100  
 Db 241 CTCACCTTACCTGCGCCACCAATGCCAGCTGTCTTCGACAGAAATCCAGAGGTGGCAG 300  
 Oy 101 GlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120  
 Db 301 GGCTACGTGCTCATGCTGCTACACCAAGTGAAGGCGAGTCCCTCCAGAGGCTGGCGATT 360  
 Oy 121 ValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAspAsnIysAsp 140  
 Db 361 GTGCGAGGACACCCAGCTCTTTGAGGACAATGCTGCCCTGGCGCTGAGCAATGAGAGAC 420  
 Oy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160  
 Db 421 CCGGTGAACAATACCACCCCTGTCCAGGGGCTCCCGAGAGGCTCCGGAGAGTGGCAG 480  
 Oy 161 LeuArgSerLeuThrGluIleLeuIysGlyValLeuIleGlnArgAsnProGlnLeu 180  
 Db 481 CTTCGAACCTTCACAGACATCTGAAAGAGAGGGGTCTTGATCCAGCGAACCCCAAGCTC 540  
 Oy 181 CysTrpTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200  
 Db 541 TGTCTACCGAGACACATTTTGTGAAAGACATCTTCACAAAGAACACAGCTGCTCTC 600  
 Oy 201 ThrIleuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220  
 Db 601 ACACGTGATGACACACCAACCGCTCTCGGGCTGTCCACCCCTGTCTCCATGTGTAAAGGC 660  
 Oy 221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240  
 Db 661 TCCCGCTCTGTGGGAGAGATTGTGAGATTGTGCAGACCTGACCGGACAGTGTCTGTCC 720  
 Oy 241 GlyGlyCysAlaValArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260  
 Db 721 GTGCGCTGTGCCCTGTCCAGGGGCTGACCTGCCACTGACTGCTGCCATGAGACAGTGTCT 780  
 Oy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280  
 Db 781 GCGGCTCAGAGGGGCCCAACAGCAGCTGACTGCTGGCTGCTGCTCCACTTCAACACAGT 840  
 Oy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGluSer 300  
 Db 841 GGCATCTGTGAGCTGCTGACCTGCCAGCCCTGTGCACCTAACACAGACAGCTTGTAGTCC 900  
 Oy 301 MetProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTrp 320  
 Db 901 ATGCCCAATCCCGAGGGCGGTATACATTCGGCGCACCTGTGTACGTCTGTCCAC 960  
 Oy 321 AsnTrpLeuSerThrAspValAlaIysSer-----GlyAla 331  
 Db 961 AACATCACTTCTACGAGAGCTGGATCCGCTAGCATCATTAATTTCGAGAACTTGGGCGCT 1020  
 Oy 332 GlyLeuMetValHisHisArgHisArgSerSerThrArgSerGlyGlyIysAspLeu 351  
 Db 1021 GGGGGCATGTGTCCACACAGGCGACGCTCATCTACAGAGAGTGGGGGAGCTG 1080  
 Oy 352 ThrLeuGlyLeuGluProSerGluGluGluAlaProAlaSerProLeuAlaProSerGlu 371  
 Db 1081 ACAGTAGGGCTGAGGCTCTGTGAAGAGAGGCCCCACAGTGTCCACTGGCACCTCCCAA 1140  
 Oy 372 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaIysGluGln 391

Db 1141 GGGGCTGGCTCCGATGATTTGATGTGACCTGGGAATGGGGGCGACCAAGGGGCTGC 1200  
 Oy 392 SerLeuProThrHisAspProSerProLeuGlnArgTrpSerGluAspProThrValPro 411  
 Db 1201 AGCTTCCACACATGACACCCCGCTCTACAGCGGTACAGTGAAGAGACCCCACTACCC 1260  
 Oy 412 LeuProSerGluThrAspGlyTrpValAlaProLeuThrCysSerProGlnProGluTrp 431  
 Db 1261 CTGGCTCTGAGACATGATGGCTACCTGCTCCCTGACTGTGACGCGCCAGCTGAATAT 1320  
 Oy 432 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 451  
 Db 1321 GTGAACACAGCGATGTGGGCCAGGCCCTTGGCCCGGAGAGGGCGCTGTGCTGCT 1380  
 Oy 452 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 471  
 Db 1381 GCCCAGCTGCTGCTGGCCACTGTGGAAAGGGCCAAAGACTCTCTCCCAAGGAAATGG 1440  
 Oy 472 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTrpLeuThrPro 491  
 Db 1441 GTCTCAAGAGCTTTTGGCTTTGGGGGTGGCGTGAAGACCCCGAGTACTGACACCC 1500  
 Oy 492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511  
 Db 1501 CAGGAGAGAGCTGCCCTCAGGCCACCCCTCTCTGCTTCAGGCCAGCTTCGACACAC 1560  
 Oy 512 LeuTrpTrpTrpAspGluAspProProGluArgGlyAlaProProSerThrPheLysGly 531  
 Db 1561 CTCTATTACTGGGACACAGACCCACACAGCGGGGCTCCACCCAGACCTTCAAAAGG 1620  
 Oy 532 ThrProThrAlaGluAsnProGluTrpLeuGlyLeuAspValProAlaAlaProAla 551  
 Db 1621 ACACCTACGGAGAGAACCCAGAGTACCTGGGTGTGAGAGTGTCCAGCGCGGCCACCC 1680  
 Oy 552 ArgSerProSerProSerThrGlnProThrGlnHisValAsnAlaIleGlnGluLysArg 571  
 Db 1681 CGCTACCCAAACCCGTGACCCCGGCTGTGAAGCATGTGATGATCAAGAAAGACTGTG 1740  
 Oy 572 ArgLeuLeuAsnLeuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValIle 591  
 Db 1741 AGCTCCCAATATGACATCGTGTGTGAGAAACGAAGAAAGAAAGAGCTGATCATTC 1800  
 Oy 592 SerGluMetPheAspLeuGlnGluProThrCysLeuGlnIleThrArgLeuGluLeuTrpLys 611  
 Db 1801 TCTAATGAGTTCTCCATCCAGAGGCGGACATGTGTGACAGCCCGCTGAACTGTATACAG 1860  
 Oy 612 GlnGlyLeuArgGlySerLeuThrLysLeuIysGlyProLeuThrMetMetAlaSerHis 631  
 Db 1861 CAGGCTTACGGGGGCAACCTCAACCAACTCAATGGCGCTTGACCATGATAGCCAGCCAC 1920  
 Oy 632 TyrLysGlnHisCysProProThrProGluThrSerCysAlaThrGlnIleIleThrPhe 651  
 Db 1921 TACAGAGCAAGTGGCTTCCAAACCCCGGAAACTGACTGAAATATGAAATGACACACCTT 1980  
 Oy 652 GluSerPheLysGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpLeu 671  
 Db 1981 GAGGATTTCAATAAAGAACTTAAGGCTTGTGTGATATCCCTTTGATGCTGTGGAG 2040  
 Oy 672 ProValGlnGluGlyAlaProProProProAlaAlaHisHisHisHisHis 690  
 Db 2041 CCGGTCCGAAGAGGCGGCCACCCCGCGG-----GGCATCAACATACACATCAC 2091  
 RESULT 3  
 AAD21564  
 ID AAD21564 standard; DNA; 1665 BP.  
 AC AAD21564;  
 XX 28-JAN-2002 (first entry)  
 DT Human HER500 fusion DNA construct.  
 DE  
 XX



Accession	Gene	Protein	Sequence	Length
QY 401	LeuGlyArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal	420		420
Db 1201	CTACAGCGGTAACACTGGAGAGCCACAGTACCTCTCCCTGTGAGACTGATGGCTACGTT	1260		1260
QY 421	AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln	440		440
Db 1261	GCCTCCCTGACCTCAGACCCCGAGCCGTAATATGTGAACACACAGATGTTGGGCCAG	1320		1320
QY 441	ProProSerProAlaGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu	460		460
Db 1321	CCCCCTTCGCCCGAGAGGGCCCTCTGCCCTGCGCCCGACCTGCTGGCCACTCTGGAA	1380		1380
QY 461	ArgAlaIaIaYrThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly	480		480
Db 1381	AGGGCCAGAACTCTCTCCCGAGAGAAATGGGGTGTGTAAAGACGTTTGTGCTTGGG	1440		1440
QY 481	GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHis	500		500
Db 1441	GGTCCCGTGGAGAACCCCGAGTACTTGACACCCACAGGAGAGCGTGGCCCTCAGCCAC	1500		1500
QY 501	ProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTTPAspGlnAspProPro	520		520
Db 1501	CTCTCTCTGCTGCTTCACCCACGCTTCGACACCTCTATTACTGGAGACAGACCCACA	1560		1560
QY 521	GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr	540		540
Db 1561	GAGCGGGGGGCTCCACCCACACCTTCAAAAGGAGACACTACGAGAGAGAACCCAGAGTAC	1620		1620
QY 541	LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSer	557		557
Db 1621	CTGGGCTCTGGACGTGCCAGCGGCC-----GCACA-TCACATCACCATCA	1664		1664
RESULT 4				
ADAD21566	standard; DNA; 1692 BP.			
ADAD21566	28-JAN-2002 (first entry)			
DE XX	Human HER500 fusion DNA construct comprising OVA-derived octapeptide.			
KW KW	Immunostimulatory fusion protein; IFP; antigen component; therapy; immunostimulatory component; T-cell mediated immune response; DC; dendritic cell; colon cancer; breast carcinoma; ovarian cancer; PAP protein; Ala Arg linker; membrane distal extracellular domain; membrane distal intracellular domain; C-terminal tag; human; OVA; HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion DNA; ds.			
OS OS	Chimeric - Homo sapiens.			
XX XX	Chimeric - Unidentified.			
FT FT	Key			
CD	Location/Qualifiers			
FT FT	1..1692			
FT FT	/*tag= a			
FT FT	/product= "Human HER500 fusion protein construct comprising human PAP signal sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal sequence, an Ala Arg linker, membrane distal extracellular domain, an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2 membrane distal intracellular domain and a C-terminal tag"			
FT FT	/note= "CDS does not include stop codon"			
FT FT	/partial			
XX XX	MO200174855-A2.			
XX XX	11-OCT-2001.			
XX XX	30-MAR-2001: 2001WO-US10515.			
XX XX	30-MAR-2000: 2000US-193504P.			

XX	(DEND-) DENDREON CORP.
PA	
XX	
PI	Laus R, Vidovic D, Gradis T;
XX	
DR	WPI: 2001-662965/76.
XX	P-PSDB: AAEL3110.
PT	An immunostimulatory fusion protein comprising the intracellular domain
PT	of HER-2 and an antigen elicits an immune response to the antigen and
PT	is useful for the treatment of associated cancer associated -
XX	
PS	Example 3: Page 28; 59pp; English.
XX	
CC	The invention relates to immunostimulatory fusion proteins (IFP) and
CC	nucleic acid molecules encoding such proteins. The IFPs comprise a
CC	polypeptide antigen component and an immunostimulatory component derive
CC	from the intracellular domain of HER-2 protein which is effective to
CC	elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC	immune response to the antigen. IFP or superactivated dendritic cells
CC	are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC	associated with a particularly antigen. The present sequence is HER500
CC	fusion DNA construct which comprises DNA molecules encoding human PAP
CC	signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
CC	signal sequence, mature HER-2 membrane distal extracellular domain,
CC	an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
CC	HER-2 membrane distal intracellular domain and a C-terminal tag.
XX	
SQ	Sequence 1692 BP: 338 A; 573 C; 470 G; 311 T; 0 other:
	Alignment Scores:
	Pred. No.: 9 3e-122 Length: 1692
	Score: 2959.50 Matches: 554
	Percent Similarity: 97.88% Conservatve: 0
	Best Local Similarity: 97.88% Mismatches: 1
	Query Match: 79.15% Indels: 12
	DB: Gaps: 2
US-09-821-883-2 (1-690) x AAD21566 (1-1692)	
OY	1 MetatgaataaiaProleuLeuLeuA1aargA1aalaSerLeuSerLeuGlyPheLeuPhe 20
DB	1 ATGAAGAGCTGCACCCCTCTCCTGGCCAGGCGAGCAACCTTAAAGCTTCTTGT 60
OY	21 LeuLeuPhePhePhePhePheLeuAspArgSerValLeuAlaArgGluLeuAlaArgGlyAlaAla 40
DB	61 CTGCTTTTCTTCTGGCTGAGACCGAAGTGTACTAGCCAAAGAGTGGCCGCGGGCCGCG 120
OY	41 SerThrGlnValCysThrGlyThrAspMetCysLeuAArgLeuProAlaSerProGluThr 60
DB	121 TCGACCCAAAGTGTGACGGCGACAGACATGAAGCTGGGGCTCCCGCCAGTCCGAGACC 180
OY	61 HisLeuAspMetLeuArgHisLeuTyGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
DB	181 CACCGGACATGCTCCGCGACCTCTACACAGGGCTGCCAGGTGTGCACAGGAACCTGGAA 240
OY	81 LeuThrTyGlnLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100
DB	241 CTCACCTACCTGCCACCAATGCCAGCTGTCTTCTCTGCACAGATATCCAGAGGTGCAG 300
OY	101 GlyTyValLeuLeuIleAlaHisAsnGlnValAArgGlnValProLeuGlnArgLeuArgIle 120
DB	301 GGGTACGGCTCATTCGCTCACAACCAAGTGAAGGCGAGTCCACATGCACAGAGCTGGGAT 360
OY	121 ValArgGlyThrGlnLeuPheGluAspAsnTyraAlaLeuAlaValLeuAspAsnGlyAsp 140
DB	361 GTGGAGGACACCCAGCTCTTGTAGGACAACTATGCGCTCGCGCTTGAACAATGAGAAC 420
OY	141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgIleuGln 160
DB	421 CCGGTGAACATATACACCCCTCTGTCAACAGGGCTCTCCACAGAGGCTTCGGGACTGTGAG 480
OY	161 LeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180

```

Db 481 CTTCAACCCCTCACAGATCTTTGAAGAGGGGCTTGATCCAGGAAACCCACAGCTC 540
Oy 181 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAspGluLeuAlaLeu 200
Db 541 TGTATCCAGGACACATTTTGTGGAAAGACATCTTCCAAAGAACACAGAGTGGCTCTC 600
Oy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACACGTGATAGAACACACCGGCTCTCGGGCTCCACCCCTGTTCCGATGTGTAAAGGC 660
Oy 221 SerArgCysTTPGlyGlnSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTCTGGGAGAGATTCTGAGATTCTCAGAGCTGACCCGACCTGTCTGTGC 720
Oy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
Db 721 GGTGGCTGTCCCGCTCCAGAGGGGCTCCACTGCTGCTGCTGCTGCTGCTGCTGCT 780
Oy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGGCTCAGCGGGGCCACAGACATCTGACTGCTGGCTGCTGCTGCTGCTGCTGCTGCT 840
Oy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlnSer 300
Db 841 GGCATCTGTGAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Oy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
Db 901 ATGCCCAATCCGAGGGCGGTATACATTCGGCGCAGCTGTGACTGCTGCTGCTGCTGCTGCT 960
Oy 321 AsnTyrLeuSerThrAspValGlySer-----GlyAla 331
Db 961 AACTACCTTTTCTAGGAGAGTGGATCCCTACATCATTAATTTCGAGAAAGTTGGCGCT 1020
Oy 332 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyLysAspLeu 351
Db 1021 GGGGCGATGTGCTCCACAGCAGCCGACGCTCATCTACAGAGAGTGGCGTGGGACCTG 1080
Oy 352 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 371
Db 1081 ACACTAGGGCTGAGACCTCTGTAAAGAGAGGCCCGGAGTCTCCACTGGCACCTCCGAA 1140
Oy 372 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 391
Db 1141 GGGGCTGGCTCCGATGATTTGATGTGACTGGGAATGGGGGACCAAGGGGCTGCAA 1200
Oy 392 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 411
Db 1201 AGCTTCCCCACACATGACCCCGCTCTACAGCGGTACAGTGAAGAGCCCGACAGTACCC 1260
Oy 412 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 431
Db 1261 CTGGCCCTGTGACATGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Oy 432 ValaGlnProAspValAlaArgProGlnProProSerProArgGlyGlyProLeuProAla 451
Db 1321 GTGAACACAGCCAGATGTGGGCCCGCCCTTGGCCCGAGAGGGCCCTTGGCTGCT 1380
Oy 452 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 471
Db 1381 GCCGCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Oy 472 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrPro 491
Db 1441 GTCTGCAAGAGCTTTTGTGCTTGTGGGGTGGCTGGAGAACCCGAGTACTTGCACCC 1500
Oy 492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511
Db 1501 CAGGAGAGAGCTGGCCCTCAGCCGACCTCTCTGCTTACGCCAGCTTTCGACAAAC 1560
Oy 512 LeuTyrTyrTTPAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGly 531

```

```

Db 1561 CTCTATTACTGGACACAGACCACAGAGCGGGGCTCCACCCAGCAGCTTCAAGG 1620
Oy 532 ThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaProAla 551
Db 1621 ACACCTACGGCAGAGAACCCAGATGACTGGTCTGGACGTGCCAGGGCC-----GCA 1674
Oy 552 ArgSerProSerProSer 557
Db 1675 CA-TCCACATGACATCA 1691

```

## RESULT 5

AA01585 standard; DNA; 4530 BP.

AA01585;

20-APR-1996 (first entry)

Her-2/neu (ERBB2/c-erbB-2) gene sequence.

Her-2/neu; Erb-B2; c-erbB-2; oncogene; DNA binding protein; HPRF;

Erb-B2 promoter binding protein; tumour enhancer factor;

breast cancer diagnosis; prognosis; antisense oligonucleotide;

retro virus vector; gene therapy vector; ss.

Homo sapiens.

W09528485-A1.

26-OCT-1995.

19-APR-1995; 95WO-US04953.

19-APR-1994; 94US-0229515.

(USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

Razluddin F, Sarkar FH.

WPI: 1995-373800/48.

New purified protein binding to the ERBB2 gene promoter - to induce

cell proliferation, diagnostic of breast cancer, also related

antibodies, nucleic acid, assays and methods for screening

inhibitors.

Disclosure: Page 52-54; 69pp; English.

The Erb-B2 gene is one of the primary genes responsible for the

transition of normal breast epithelial cells towards carcinoma in

situ and the subsequent development of invasive and metastatic

cancer. HPRF (see AAR77093-94), the Erb-B2 promoter binding protein,

induces cell division on binding to the promoter. In a method for

greater success in early identification and treatment of breast

cancer, the initiation step for Erb-B2 gene activity is identified.

This method involves determining the presence of HPRF in a biopsy

from the subject, where the presence of HPRF (relative to its

absence in a normal control) indicates the presence of cancer and

a decreased chance of long-term survival. Binding of HPRF to the

CC promoter can be inhibited using antisense oligonucleotides or a

non-genomic nucleic acid that binds to HPRF; these oligos can be

expressed from retro virus or other gene therapy vectors.

Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other;

## Alignment Scores:

Pred. No.: 8,15e-98 Length: 4530

Score: 2423.00 Matches: 524

Percent Similarity: 41.14% Conservative: 5

Best Local Similarity: 40.75% Mismatches: 21

Query Match: 64.80% Indels: 736

DB: 16 Gaps: 4





Oy	329	-----	329	-----	329
Db	2245	CCGCTGACACCTTACGGAGCGGATGCCCAACACAGCGCAGATCGGATCCTGAAGACAGC	2304		
Oy	329	-----	329	-----	329
Db	2305	GAGCTGAGGAAGTGTAAGTGCTTGGATCTGGGCGCTTTTGGCACAGTCTACAGGGCATC	2384		
Oy	329	-----	329	-----	329
Db	2365	TGGATCCCTGATGGGAGAGATGTGAAAATTCCAGTGGCCATCAAGTGTTAGGGAAAC	2424		
Oy	329	-----	329	-----	329
Db	2425	ACATCCCCCAAGCCACAAAGAAATCTTAGACGAAGCATACGTAGTGCTGTGTGGC	2484		
Oy	329	-----	329	-----	329
Db	2485	TCCCCATATGTCTCCGCGCTTTGGGCATCTGCCTACATCCACGGTGACGTGTGTACA	2544		
Oy	329	-----	329	-----	329
Db	2545	CAGCTATGCCCTATGGCTGCTCCTTTAGACCATGTCCGGGAAACCGGGAGCGCTGGGC	2604		
Oy	329	-----	329	-----	329
Db	2605	TCCAGAGACCTGTAAGTGTGTATGCAGATTGCCAAGGGATGACCTAGCTGAGAGAT	2664		
Oy	329	-----	329	-----	329
Db	2665	GTGGGGCTCGTACACAGGGACTTGGCGCGCTCGAAGCTGCTGTCAAGATCCACCAT	2724		
Oy	329	-----	329	-----	329
Db	2725	GTCAAAATTACAGACTTCGGGCTGGCTCGGCTGCTGACATTGACAGACAGAGTACCAT	2784		
Oy	329	-----	329	-----	329
Db	2785	GCAGATGGGGGGAAGTGCCCATCAAGTGATGGCGCTGAGTCCATTCTCCGCGCGG	2844		
Oy	329	-----	329	-----	329
Db	2845	TTGACCCACCAAGAGTGATGTGTGAGATTATGTTGCTACTGTGGAGCTGATGACTTT	2904		
Oy	329	-----	329	-----	329
Db	2905	GGGGCCAAACCTTACGATGGGATGCCAGCCGGAGATCCTGACCTGCTGAAAAGGG	2964		
Oy	329	-----	329	-----	329
Db	2965	GAGCGGCTGCCCCAGCCCCCATCTGCACCATTTGATCTACATGATCATGTGTCAAATGT	3024		
Oy	329	-----	329	-----	329
Db	3025	TGGATGATTACTGTGAATGTGGCCCAAGATTCGGGAATTGGTGTGAATTTCTCCGC	3084		
Oy	329	-----	329	-----	329
Db	3085	ATGGCCAGGAGCCCCCAGCGCTTTGTGTGCATCCAGAAATGAGAGACTTGGGCCACCGACT	3144		
Oy	329	-----	329	-----	329
Db	3145	CCCTTGGACAGCACCTTCTACCGCTCAGCTGAGAGAGATGACATGGGGACCTGTGTG	3204		
Oy	330	-----	330	-----	330
Db	3205	GATGCTGAGAGATATCTGTGTAACCCACAGCAGGCTTTTCTGTCCAGACCCCTGCCCCGGC	3264		
Oy	331	AlaGlyIyMetValIhIshIstArgHisArgSerSerSerThrArgSerGlyIyIyIyAsp	350		
Db	3265	GCTGGGGGCAATGGTCCACACAGGCAACCGAGCTCATCTACACAGGAGTGGCGTGGGAC	3324		
Oy	331	LeuThrIeuGlyIeuGluProSerGluGluAlaProArgSerProLeuAlaProSer	370		

Db	3325	CTGACACTAAGGGCTGGAGCCCTCTGGAAAGAGAGGCCCCCAGGCTCCACTGACCTCC	3384
Qy	371	GIUGIYALAGISerAspValAPheAspGIYAspleuGIYmetGIYAlaAlaLYsgIYleu	390
Db	3385	GAAGGGCGCTGGCTCCGATGATTTTGATGATGACCTGGGAATGGGGGAGGCAAGGGGCTG	3444
Qy	331	GInserLeuProThrHisAspProSerProLeuGInaYrSerGIYAspProThrVal	410
Db	3445	CAAAACCTCCCCACACATAGACCCACGCTCTACAGCGGTACATGAGGAGCCACAGTA	3504
Qy	411	ProLeuProSerGIYThrAspGIYTYValAlaProLeuThrcysSerProGInProGlu	430
Db	3505	CCCCGCCCCCTGAGACTGATGCTACGTGCCCCCTGACTGACAGCCCCAGCCTGAA	3564
Qy	431	TYrValaEngInProAspValAlaProGInProProSerProArgGIYGIYProLeuPro	450
Db	3565	TATGTGAACCAAGCCAGATGTTGGGCCCCAGCCCTTCGCCCGAGAGGGCCCTGTGCT	3624
Qy	451	AlaAlaArGProAlaGIYAlaThrLeuGInaYrAlaLYsThrLeuSerProGIYLYsAsn	470
Db	3625	GCTGCCCAACCTGCTGGGTGGCCACTGCTGGAAAGGGCCAAAGACTCTCTCCAGAGGAAT	3684
Qy	471	GIYValIYValLYsAspValAPheAlAPheGIYGIYAlaValGIYAsnProGIYTYrLeuThr	490
Db	3685	GGGGTCGTCAAAGAGGTTTTGGCTTTGGGGTCCCTGTGGAGAACCCCGAGTACTTGACA	3744
Qy	491	ProGInGIYGIYAlaAlaProGInProHisProProProAlaPheSerProAlaPheAsp	510
Db	3745	CCCCAGGAGAGAGCTGCCCTCAGCCCCAGCCCTCTCTGCTTCAAGCCAGCTTGAC	3804
Qy	511	AsnLeuTYrTYrTPAspGInAspProProGInaYrGIYAlaProProSerThrPheLYs	530
Db	3805	AACCTCTATTACTGGGACCAAGACCCACAGAGCGGGGGCTCCACACCACTTCATAA	3864
Qy	531	GIYThrProThrAlaGInaAsnProGInTYrLeuGInLYsAspValAProAlaAlaAlaPro	550
Db	3865	GGGACACTACGGCAGAGAACCCAGAGTACTGGGTGGACGTGCCA-----GTGTGA	3918
Qy	551	AlaArGSerProSerPro 556	
Db	3919	ACCCAGAGCCCAAGTCCG 3936	
RESULT 6			
ID	AAAT71253	AAAT71253 standard; DNA; 4530 BP.	
XX	AAAT71253;		
AC	AAAT71253;		
XX			
DT	30-MAR-1998 (first entry)		
XX			
DE	Human HER2 gene.		
XX			
KW	HER2; cognate transgene; human; tyrosine kinase -type receptor; Lymphoma;		
KW	cellular immunogen; cancer; self-determinant Immunoreactivity;		
KW	Cancer vaccination; breast carcinoma; colon carcinoma; Immunotherapy;		
KW	proto-oncogene; ss.		
OS	Homo sapiens.		
XX			
PN	W09725860-A1.		
XX			
PD	24-JUL-1997.		
XX			
PF	13-JAN-1997; 97WO-US00582.		
XX			
PR	19-JAN-1996; 96US-0010262.		
XX			
PA	(UYAL-) UNIV ALLEGHENY HEALTH SCI.		
XX			
PI	England JM, Halpern MS;		
XX			
DR	WPI; 1997-384993/35.		



329  
Db 1765 GAGTGGCGGAGGAGTCCGAGTACTGACGAGGGCTCCCCAGGAGTATGTGAATGCCAGG 1824  
329  
Db 1825 CACTGTTGGCGTGACACCTGAGTGTCAAGCCCCGAAATGAGTGCAGTACCTGTTTGA 1884  
329  
Db 1885 CCGGAGGCTGACAGTGTGTGGCTGTGCCACTATAAGAGACCTCTCTGCGTGCC 1944  
329  
Db 1945 CGCTGCCAGCGGTGTAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2004  
329  
Db 2005 GAGGAGGCGCATGCGACGCTTGGCCCATCAACTGACACCACTCCTGTGTGGACCTGGAT 2064  
329  
Db 2065 GACAAGGCTGCCCGCCGAGCAGAGACGACCCCTGTGACGTCCATGCTCTGCGGTG 2124  
329  
Db 2125 GTTGGCATTTCTGCTGCTGTGCTCTTGGGGGTGGTCTTTGGGATCTTCATCAAGCGACGG 2184  
329  
Db 2185 CAGCAGAGATCCGGAAGTACAGCATGCGGAGACTGTGCAAGAAAGGAGCTGTGTAG 2244  
329  
Db 2245 CCGCTGACACCTAGCGGAGCGATGCCAACAGCGGAGATGCCGATCTGAAAGAGAG 2304  
329  
Db 2305 GAGCTGAGGAGGTGAAGTGTGATCTGCGCTTTTGGCACAGTCTACAGGGCATC 2364  
329  
Db 2365 TGGATCCCTGATGGGAGATGTGAAATTCAGTGGCCATCAAGTGTGAGGGAAC 2424  
329  
Db 2425 ACATCCCCCAAGCCAAACAAAGAAATCTTGAAGCAAGCATACGTGATGGCTGTGGC 2484  
329  
Db 2485 TCCCCATATGTCTCCCGCTTCTGGGCATCTGCGTACATCCAGCGTCCAGCTGTGACA 2544  
329  
Db 2545 CAGCTTATGCCCTATGCTGCGCTTTAGACCATGTCCGGAACCGCGGAGCGCTGGGC 2604  
329  
Db 2605 TCCGAGACCTGCTGAAGTGTATGAGATTGCCAAGGGGAGTACCTACCTGAGGAT 2664  
329  
Db 2665 GTGGGGCTCGTACACAGGAGCTTGGCCGCTCGGAACGTGTGTCAAGTCCCAACAT 2724  
329  
Db 2725 GTCAAAATTAAGACTTGGGCTGGCTGGCTGTGACATTTGACAGACAGATACCAT 2784  
329  
Db 2785 GCAGATGGGGGCAAGGTCCCATCAAGTGAATGCGCTGAGTCCATTCTCCGCGCGG 2844

329  
Db 2845 TTCACCCACAGAGTGAATGTGTGAGATTATGTGTGACTGTGTGGAGACTGATGACTTTT 2904  
329  
Db 2905 GGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGTGAAAAGGGG 2964  
329  
Db 2965 GAGCGCTGCCACAGCCCCCATCTGACCATTTGATGTCTACATGATCATGTGTAATGT 3024  
329  
Db 3025 TGGATGATTGACTGTGAATGTGCGCAAGATTCCGGAGTGTGTCTGAATTTCTCCGC 3084  
329  
Db 3085 ATGGCCAGGAGACCCACGCTTTGTGTGATCCAGAAATGAGAGATGGCCACAGCAT 3144  
329  
Db 3145 CCGTTGACACACCTTCTACCGCTCACTGCTGAGAGCATGACATGGGGACCTGTG 3204  
330  
Db 3205 GATGCTGAGAGTATCTGTGATCCAGAGGGCTTCTGTCCAGACCCTGCCCGGGC 3264  
331  
Db 331 AlaGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyIleVal 350  
3265  
Db 3265 GCTGGGGCATGTGTCACACAGGACCGACGACATCATCTACAGAGATGGCGGTGGGAC 3324  
351  
Db 351 LeuThrLeuGlyLeuGluProSerGluGluIleAlaProAlaGlySerProLeuAlaProSer 370  
3325  
Db 3325 CTGACACTAGGGCTGGACCCCTGTGAAGAGAGGCCCCAGGTCTCCACTGGACCTCC 3384  
371  
Db 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlySerIleu 390  
3385  
Db 3385 GAAGGGCTGGCTCCGATGTATTTGATGTGATGATGATGATGATGATGATGATGATGATGAT 3444  
391  
Db 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410  
3445  
Db 3445 CAAGCCCTCCACACATGACAGCCCGCTCTACAGCGGTACAGAGAGAGAGAGAGAGAGAGAG 3504  
411  
Db 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430  
3505  
Db 3505 CCGCTGCCCTGTGAGACTGATGTGCTACGTGCCCCCTGACCTGACCTGACCCCGAGCTGAA 3564  
431  
Db 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 450  
3565  
Db 3565 TATGTGAACACAGCCAGATGTTGGGCCAGCCCTTGGCCCGAGAGGGCCCTTGCTGCT 3624  
451  
Db 451 AlaAlaArgProAlaGlyAlaThrLeuGluAlaValGlyIleValGluAspProGluIleuThr 470  
3625  
Db 3625 GCTGCCACCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3684  
471  
Db 471 GlyValIleValAspValPheAlaPheGlyIleValGlyIleValGlyIleValGlyIleVal 490  
3665  
Db 3665 GGGGTGTCAAAGAGCTTTTGGCTTGTGGGGTGGCTGGAGAAACCCGAGTACTGTGACA 3744  
491  
Db 491 ProGlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp 510  
3745  
Db 3745 CCGCAGGAGAGAGCTGCCCTCAGCCACCTCTCTGCTTCCAGCCACCTTTCAGC 3804  
511  
Db 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheIys 530  
3805  
Db 3805 AACCTCTATTACTGGGACAGAGACCCAGAGCGGGGGGCTCCACCCAGCACCTTCAA 3864  
531  
Db 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaIlePro 550  
3865  
Db 3865 GGGACACTAGGACAGAAACCCAGATACCTGGGTCTGGAGCTGCA-----GTGTGA 3918  
551  
Db 551 AlaArgSerProSerPro 556

```
DB 3919 ACCAGAGGCCAAGTCCG 3936
|||||
RESULT 7
AAZ60815
ID AAZ60815 standard; DNA; 4530 BP.
XX
AC AAZ60815;
XX
DT 16-MAY-2000 (first entry)
XX
DE Nucleotide sequence of a cognate transgene of c-neu..
XX
KW Cognate transgene; CTG; tumourigenic; cellular immunogen; immunisation;
KM proto-oncogene; malignancy; allogenic cell; vaccine; cancer; ss.
XX
OS Homo sapiens.
XX
PN W0200004927-A1.
XX
PD 03-FEB-2000.
XX
PF 08-JUL-1999; 99MO-US15594.
XX
PR 24-JUL-1998; 98US-0093965.
XX
PA (UYAL-) UNIV ALLEGHNEY HEALTH SCI.
PA (HALP/) HALPERN M S.
PA (ENGL/) ENGLAND J M.
XX
PI Halpern MS, England JM;
XX
DR WPI; 2000-182543/16.
XX
PT Cellular immunogens comprising allogenic donor cells transfected with a
PT construct comprising a proto-oncogene cognate, useful as cancer
PT vaccines -
XX
PS Disclosure: Page 66-68; 77pp; English.
XX
CC The present sequence represents a cognate transgene (CTG) which is
CC rendered non-tumourigenic by deletion of amino acids 1-731. The CTG
CC is used in the course of the invention. The specification describes a
CC cellular immunogen for immunizing a host against the effects of the
CC product of a target proto-oncogene which is associated with a
CC malignancy. The cellular immunogen comprises allogenic cells
CC transfected with transgene construct comprising a transgene cognate
CC to target proto-oncogene and a strong promoter. The cellular immunogen
CC is useful for vaccinating a host against cancer by inserting the
CC transgene construct into the body of the host for the expression of
CC the transgene. The method of the invention is designed to target
CC mutation-driven non-self determinants. The cellular immunogens induce
CC reactivity for self-determinants in the over expressed product of tumour
CC associated and over expressed proto-oncogenes.
XX
SQ Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other;

Alignment Scores:
Pred. No.: 8,15e-98 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 21
Query Match: 64,808 Indels: 736
DB: 21 Gaps: 4

US-09-821-883-2 (1-690) x AAZ60815 (1-4530)
QY 2 ArgAlaAlaProLeuLeuLeuAlaAla-----AlaSerLeuSerLeuGly 17
DB 94 CGGCGCCCTCCCGACCGGTCACCGAGCCATGGGCGCGCCGACCGAGTACACCATG 153
QY 18 PheLeuPheLeuPhePhePheTrPLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
|||||
```

```
DB 154 GAGCTGGCGGCTTTGTCCTCCGCTG-----GGCTCTCTCTCCGCTCTTGGCCCCC 204
QY 38 GlYAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
DB 205 GGAGCCCGCGAGACCCAAAGTGTGACCGGACACAGACATGAAGCTGGGGCTCCGCGCAGT 264
QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
DB 265 CCCGAGACCCACCTGGACATGCTCCGCCACCTTCACAGGAGGCTCCAGGTGTCCAGGGA 324
QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
DB 325 AACCTGGAACTCACTACCTCTGCCCCACCAATGCCAGCTGTCTTCTTCAGAGATATCCAG 384
QY 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
DB 385 GAGGTGCGAGGCTATCGTGCATGCTGTCACACACCAAGTGAAGGCGATGCCACTGCAGAGG 444
QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
DB 445 CTGCGGATTTGCGGAGCGCACCCAGCTCTTGAGAGCAACTATGCCCTTGCGCGCTGTAGAC 504
QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlnValAspProGlnGlyLeuArg 157
DB 505 AATGAGAACCCCGCTGAACAAATACCAACCCCTGTCCAGAGGCGCTCCCGAGAGGCTCGG 564
QY 158 GluLeuGlnLeuArgSerLeuThrGlnIleLeuLysGlyLeuValLeuIleGlnArgAsn 177
DB 565 GAGCTGCAAGCTTGGAACCTTCACAGAGATCTTGAAGAAGAGGGGTCTTGATCCACGGAAAC 624
QY 178 ProGlnLeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysValAsnGln 197
DB 625 CCCGAGCTGTCTACACGAGACACAGATTTGTGAGAGACATCTTCCAAAGAACACACAG 684
QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
DB 685 CTGGCTTCACACTGATAGACACACCAACCCCTCTGGGCTCCGACCCCTGTCTCCCATG 744
QY 218 CysLysGlySerArgCysTyrGlyGlyLeuSerSerGlnAspCysGlnSerLeuThrArgThr 237
DB 745 TGTAAAGGGCTCCGCTCTGCTGGGAGAGATTTGTAGAGATTTGTAGACCTTCACCGCCACT 804
QY 238 ValCysAlaGlyGlnCysAlaArgCysAlaGlyGlyProLeuProThrArgCysCysHisGly 257
DB 805 GTCTGTCCGCTGCTGTGCTGCTGCTGCAAGGGGCGGCTGCGGCTGCTGCTGCTGCTGCTG 864
QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerArgCysLeuAlaCysLeuHisPhe 277
DB 865 CAGTGTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
QY 278 AsnHisSerGlyIleCysGlyLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
DB 925 AACCCAGATGGCATCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984
QY 298 PheGluSerMetProAsnProGlnGlyArgTyrThrPheGluAlaSerCysValThrAla 317
DB 985 TTTGAGTCCATGCCCAATCCCGAGGCGGATATACATTCGAGCGCACTGTGTGACGTGCC 1044
QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer-----?-----?----- 329
DB 1045 TGTCCCTCAACAACTATCTTACGAGAGTGGGATCTTCGACACCTCTGCTGCTGCTGCTGCT 1104
QY 329 ----- 329
DB 1105 AACCAAGAGGTGACAGCAGAGATGAGAACACAGCGGTGTGAGAAGTGCACAGACCTGT 1164
QY 329 ----- 329
DB 1165 GCCCGAGCTGTATATGCTGTGGCGATGAGACACTTGCAGAGAGTGAAGGCGATTTCCAGT 1224
QY 329 ----- 329
DB 1225 GCCAATATCCAGAGATTTGCTGCTGCTGCTGCAAGAAATCTTTGGAGGCTGGGATTTCTGCCG 1284
```

QY	329	-----	329	-----	329	-----
Dp	1285	GAGAGCTTTATGATGGGACCCAGCCTCCAAACACTGCCCGCTCCAGCCAGAGCACTCCAA	1344	-----	1344	-----
QY	329	-----	329	-----	329	-----
Dp	1345	GTGTTTGAGACTGTGGAAAGATCACAGGTTAOCCTATACATCTCAGCATGGCCGACAGC	1404	-----	1404	-----
QY	329	-----	329	-----	329	-----
Dp	1405	CTGCCTGACCTCAGCGCTCTTCCAGAACCTGCAGATATCCGGGAGCAATTCGCACAA	1464	-----	1464	-----
QY	329	-----	329	-----	329	-----
Dp	1465	GGCCCTACTCGCTGACCTTCGAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCACTG	1524	-----	1524	-----
QY	329	-----	329	-----	329	-----
Dp	1525	AGGGAACTGGGCAGTGGACTGGCCCTCATACACATATACACCACCTGCTTCGTGCAC	1584	-----	1584	-----
QY	329	-----	329	-----	329	-----
Dp	1585	ACGGTCCCTTGGAGCCAGCTCTTTCCGAACCCGACCAAGCTCTGCTCCACTGCCAAC	1644	-----	1644	-----
QY	329	-----	329	-----	329	-----
Dp	1645	CGGCGAGAGACGAGTGTGTGGGCGAGAGGCGCTGGCTCGCACAGGTGTGGCCCGAGG	1704	-----	1704	-----
QY	329	-----	329	-----	329	-----
Dp	1705	CACCTGTGGGATCCAGGGCCACCCAGCTGTCACTGACGCACTTCCTTCGGGGCCAG	1764	-----	1764	-----
QY	329	-----	329	-----	329	-----
Dp	1765	GAGTGCCTGGAGGAAATCCGAGTACTGCAAGGGCTCCCGAGGAGATGTGAATCCAGG	1824	-----	1824	-----
QY	329	-----	329	-----	329	-----
Dp	1825	CACCTGTTGCCGTGCCACCCCTGAGTGTACAGCCCAAGATGCTCAGTACCTGTTTGA	1884	-----	1884	-----
QY	329	-----	329	-----	329	-----
Dp	1885	CCGGAGCTGACCAAGTGTGTGGCCTGTGCCACTATAAGGACCTTCCTTGCCTGGCC	1944	-----	1944	-----
QY	329	-----	329	-----	329	-----
Dp	1945	CGCTGCCCAAGCGGTGTGAACCTGACCTCTCCTACATGCCATCTGGAAGTTTCCAGAT	2004	-----	2004	-----
QY	329	-----	329	-----	329	-----
Dp	2005	GAGAGGGCGCATGCCAGCCTTCCCCCAATCACTGCACCCACTCCTGTGTGAACTGGAT	2064	-----	2064	-----
QY	329	-----	329	-----	329	-----
Dp	2065	GACAAGGCGTCCCGCCCGCAGAGAGAGCGACGCCCTCAGCTCCATGTCTCTCGGATG	2124	-----	2124	-----
QY	329	-----	329	-----	329	-----
Dp	2125	GTTTGATCTGTCTGTGTGTGTCTTTGGGGGTGTCTTTGGGATCCTCATCAAGCAGCG	2184	-----	2184	-----
QY	329	-----	329	-----	329	-----
Dp	2185	CAGCAGAAGATCCGGAAGTACACGATGCGGAGACTCTCGAGAAACGAGACTGGTGAG	2244	-----	2244	-----
QY	329	-----	329	-----	329	-----
Dp	2245	CCGCTGACACCTACGGAGGATGCCCAACAGAGCGCAATGGGATCTCTGAAGAGACG	2304	-----	2304	-----
QY	329	-----	329	-----	329	-----
Dp	2305	GAGCTGAGGAAGTGTAAGTGTCTTGGATCTGGCGCTTTTGGACAGTCTACAAAGGCA	2364	-----	2364	-----

QY	329	-----	329	-----	329
Db	2365	TGAGTCCCTGATGGGGAGAAATGTGAAAATTGCAGTGGCCATCAAACTGTTAGGCGTAAAC	2424	-----	2424
QY	329	-----	329	-----	329
Db	2425	ACATCCCCCAAGCAACCAAGAAATCTTAGAGAAAGCATACGTGATGGCTGTGGGC	2488	-----	2488
QY	329	-----	329	-----	329
Db	2485	TCCCCATATGTCTCCCGCTTCTGGGCATCTGCATCCACGGGTGACCTGGTGACA	2544	-----	2544
QY	329	-----	329	-----	329
Db	2545	CAGCTTATGCCCTATGGCTGCCTTTAGACCAATGTCGGGAAAAACCGCGAGCTGGGC	2604	-----	2604
QY	329	-----	329	-----	329
Db	2605	TCCAGAGCATGCTGAACTGTGTATGCATATGCCAAGGGATGAGTACTCGGAGAT	2666	-----	2666
QY	329	-----	329	-----	329
Db	2665	GTGCGGCTGTAACACAGGAGACTGGCCGCTCGGAACGTGCTGTCAAGATCCCAACAT	2724	-----	2724
QY	329	-----	329	-----	329
Db	2725	GTCAAATTTACAGACTTCGGGCTGGCTCGGCTGTGCATTTGACGAGACAGATGACAT	2784	-----	2784
QY	329	-----	329	-----	329
Db	2785	GCAGATGGGGCAAGGTGCCATCAAGTGGCTGGAGTCCATTTCTCCGCGGGCG	2844	-----	2844
QY	329	-----	329	-----	329
Db	2845	TTCAACCACAGAGTGAATGTGGAGTTATGTGTACTGTGGAGAGCTGATGACTTT	2904	-----	2904
QY	329	-----	329	-----	329
Db	2905	GGGGCAAACTTACGATGGGATCCAGCCGGGAGATCCGTGACCTGTGAAAAAGGG	2966	-----	2966
QY	329	-----	329	-----	329
Db	2965	GAGGGCTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAATGT	3024	-----	3024
QY	329	-----	329	-----	329
Db	3025	TGGATGATTGACTCTGAATGTGGCCAAAGATTCGGGGAGTTGGTGTGAATTTCTCCGC	3084	-----	3084
QY	329	-----	329	-----	329
Db	3085	ATGGCCAGGAGACCCCAAGCGTTTGTGTCATCCAGATGAGACTTGGGCCCAAGCACT	3144	-----	3144
QY	329	-----	329	-----	329
Db	3145	CCCTTGGACAGCACTTCTACCGCTCACTGCTGGAGAGCATGACATGGGGAGCTGGTG	3204	-----	3204
QY	330	-----Gly	330	-----	330
Db	3205	GATGCTGAGGATATCTGATACCCACAGAGGCTTCTTCTGTCACAGACCCTGCCCGGC	3266	-----	3266
QY	331	AlaGlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAsp	350	-----	350
Db	3265	GCTGGGGGCAATGGTCCACACAGCACCGAGCTCACTTCCAGGAGTGGCGTGGGAC	3324	-----	3324
QY	351	LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer	370	-----	370
Db	3325	CTGACACTAAGGGCTGGAGCCCTCTGAAGAAGAGGCCCCCAGAGTCTCCACTGGGACCTCC	3388	-----	3388
QY	371	GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyValAlaIleArgGlyLeu	390	-----	390
Db	3385	GAAGGGCGTGGCTCCGATGTATTTGATGAGTGACTCTGGAAATGGGGGCAACCAAGGGCTG	3444	-----	3444
QY	391	GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal	410	-----	410

```

Db 3445 CAAAGCCTCCACACATGACCCAGCCCTCTACAGCGGTACGTGAGAGACCCACAGTA 3504
Qy 411 ProleuprosersglutthraspgrlytyrvalaAlaProleuthrcyserserproglu 430
Db 3505 CCCCTGCCCTCTGTAGACTGAGGTGCTAGCTGCCCCCTGACCTGCACCCCGACCTGAA 3564
Qy 431 TyrtValasnglnProaspvalarprogluProproserProaArgluglyProleupro 450
Db 3565 TATGTGAACGACGAGATGTTGGGCCAGCCCTTCGCCCGAGAGGGCCCTCTGCT 3624
Qy 451 AAlaAlarproAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 470
Db 3625 GCTGCCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3684
Qy 471 GLyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 490
Db 3685 GGGGTGCTCAAAAGAGCTTTTGGCTTTGGGGGTGCTGCTGCTGCTGCTGCTGCTGCT 3744
Qy 491 ProgluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 510
Db 3745 CCCCAGGAGAGAGCTGCCCTGCACCCAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCT 3804
Qy 511 AsnleutytyrtyrtpaspgrlnaspproprogluArglyAlaAlaAlaAlaAlaAla 530
Db 3805 AACCTCTATTAAGTGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3864
Qy 531 GLythProthralaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 550
Db 3865 GGGACACTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 3918
Qy 551 AlaArgserProserPro 556
Db 3919 ACCGAAAGGCCAAGTCCG 3936
RESULT 8
AADI9731
ID AADI9731 standard; cDNA; 4530 BP.
XX
AC AAD19731;
XX
DT 18-Dec-2001 (first entry)
XX
DE Human tyrosine kinase-type receptor. HER-2 cDNA.
XX
KW Therapeutic compound; major histocompatibility complex; vaccine;
KW antigenic peptide; MHC; immunoregulatory; immune response; HER-2;
KW adoptive immunotherapy; anti-cancer; breast cancer antigen; APC;
KW antigen presenting cell; human; tyrosine kinase-type receptor; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 151..3198
FT /tag=a
FT /product="Human tyrosine kinase-type receptor,
FT HER-2"
XX
PN MO20016867-A2.
XX
PD 20-SEP-2001.
XX
PE 16-MAR-2001; 2001MO-USA0328.
XX
PR 16-MAR-2000; 2000US-0527487.
XX
PA (GENZ ) GENZYME CORP.
XX
PI Nicolette CA;
XX
DR WPI: 2001-616284/71.
XX
DR P-PSDB: AAEI2130.
XX

```

```

PT Novel synthetic therapeutic compound for inducing immune response and
PT for use in adoptive immunotherapy, has enhanced binding to major
PT histocompatibility molecules and enhanced immunoregulatory properties
PT
PS Disclosure; Page 57-63; 69pp; English.
XX
XX The invention relates to synthetic therapeutic compounds (antigenic
CC peptides) with enhanced binding to major histocompatibility complex
CC (MHC) molecules and enhanced immunoregulatory properties relative
CC to their natural counterparts. Compounds of the invention are useful
CC for inducing an immune response in a subject and for use in adoptive
CC immunotherapy. They are useful as components of anti-cancer vaccines
CC and to expand immune effector cells that are specific for cancers
CC characterised by expression of the breast cancer antigen, HER-2.
CC Polynucleotides that encode peptides of the invention are useful as
CC hybridisation probes and as primers for the detection of genes of gene
CC transcripts that are expressed in antigen presenting cells (APCs), to
CC confirm transduction of polynucleotides into host cells. The present
CC sequence is human tyrosine kinase-type receptor, HER-2 cDNA.
XX
XX
SQ Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other:
XX
XX
Alignment Scores:
Pred. No.: 8.15e-98 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 21
Query Match: 64.80% Indels: 736
DB: 22 Gaps: 4
US-09-821-883-2 (1-690) x AADI9731 (1-4530)
Qy 2 ArgAlaAlaProleuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 17
Db 94 CGGCCCTCCAGCCGGGTCCAGCCGAGCCATGGGCCGAGCCGACAGTACACCATG 153
Qy 18 PheLeuPheLeuLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 37
Db 154 GAGCTGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 204
Qy 38 GLyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 57
Db 205 GGACCGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 264
Qy 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValAlaGlnGly 77
Db 265 CCGGAGACCCACCTGACATGCTCCGACCTTACACAGGGCTCCAGGTGTGACAGGA 324
Qy 78 AsnLeuGluLeuThrTyrluProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325 AACCTGAACTCACCTACCTGCCCCACCAATGCCAGCTGCTTCCTGACGATATCCAG 384
Qy 98 GLyValGlnGlyTyrluAlaAlaAlaHisAsnGlnValArgGlnValProleuGlnArg 117
Db 385 GAGGTGAGGGCTACGCTGCTATGCTCACAAAGAGGAGGCTCCACGCTCAGAGG 444
Qy 118 LeuArgIleValArgGlyThrGlnLeuPheGlnAspAsnTyrlaAlaAlaValAlaLeuAsp 137
Db 445 CTGGGATGTGGAGGACCCACCTCTTGTAGAGACAATATGCGCGCGCTGTAGAC 504
Qy 138 AsnGlyAspProleuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 505 AATGAGAGCCGCTGAAACAATACACCCCTGTACAGAGGGCTCCACGAGGAGGCTCGG 564
Qy 158 GlnLeuGlnLeuArgSerLeuThrGlnIleLeuysGlyValAlaIleGlnArgAsn 177
Db 565 GAGCTGAGCTTGAAGCTTCACAGAGATCTTGAAGAGAGGGCTCTTGTATCCAGGAAAC 624
Qy 178 ProGlnLeuCysTyrglnAspThrIleLeuThrPylAspIlePheHisIlyAsnAsnGln 197
Db 625 CCCCAGCTGCTGCTACAGAGACAGATTTTGTGAAAGAGACATCTTCCAAAGAACACACAG 684

```

OY	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet	217
Db	665	CTGGCTCTCAACAGTATAGACACCAACCGCTCTGGGGCTGGCACCCCTGTTCTCCAGT	744
OY	218	CysLysGlySerMetArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThr	237
Db	745	TGTAAAGGGCTCCGGCTCTCTGGGAGAGAGTTCTGAGAGATTGTTCAGAGCCTGACGGCACT	804
OY	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu	257
Db	805	GTCGTGCTCCGGTGGCTGTGGCCCTGCACAGGGGCCACTGCCACATGACGTGCTGCATAG	864
OY	258	GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisIle	277
Db	865	CAGTGTCTCGCGGCTGTACAGGGCCCCACAGACACTGTGACTGCTGGCTGCTCCACTTC	924
OY	278	AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297
Db	925	AACCAACGTGGCACTGTGAGCTGCACCTGGCCAGCCCTGTGTAACCTACACACACACAGC	984
OY	298	PheGlnSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCysValThrAla	317
Db	985	TTTGAGTTCATGCCCAATCCCGAGGGCGGTATATACATCTGGGGCCAGCTGTGACTGCC	1044
OY	318	CysProTyrIleAsnTyrLeuSerThrAspValGlySer-----	329
Db	1045	TGTCCCTACAACTACTCTTTCTACGGAGCTGGGATCCTGCACCCCTGTCTGCCCTGCAC	1104
OY	329	-----	329
Db	1105	AACCAAGAGTGCACAGCAGAGATGGAAACACAGCGGTGTGAGAAAGTGCACAGACCCCTGT	1164
OY	329	-----	329
Db	1165	GCCGAGTGTGCTATGTGTCTGGGATATGAGACACTTGCAGAGGTGAGGGCAATTACAGT	1224
OY	329	-----	329
Db	1225	GCCAAATACAGAGATTGTGTGGCTGCACAGAAAGATCTTTGGAGAGCTGTGCTGCGG	1284
OY	329	-----	329
Db	1285	GAGAGCTTTGATGGGGGACCAAGCTCCACACTGCCCGCTCCAGCCACAGCAGCTCCAA	1344
OY	329	-----	329
Db	1345	GTTGTTGAGACTCTGGAAGAGATCACAGGTTACCTATTACATCTCAGCATGGCGGGACAGC	1404
OY	329	-----	329
Db	1405	CTGGCTGACCTCAGCGTCTTCCAGAAAGCTTCAGATTAATCCGGGAGAGAAATTCTGCACAT	1464
OY	329	-----	329
Db	1465	GGCGCCTACTCGCTGACCTTCGAAGGGGTGGGCAATCAGTGGCTGGGGCTGGCTCACTG	1524
OY	329	-----	329
Db	1525	AGGGAAGTGGGCAAGTGGAGCTGGCCCTCATCCACCAATAAACCCACACTTGTCTGGTGAC	1584
OY	329	-----	329
Db	1585	ACGGTCCCTGGGACACAGCTTTTCGGAACCGGCACACCAAGCTCTGTGCTCACACTGCCAAC	1644
OY	329	-----	329
Db	1645	CGGCCAGAGACGAGTGTGTGGCGAGGGCTGGCTGCCACAGCTGTGCGCCGAGGG	1704
OY	329	-----	329
Db	1705	CACGTCTGGGGTCCAGGGCCCCACCAAGTGTGTCACTGCAGCCAGTTCCTTGGGGCCAG	1764
OY	329	-----	329

Db	1765	GAGTCCGTGGAGGAATGCCAGTACTGACAGGGGCTCCCAAGGAGATGTGAATGCCAGG	1824
Qy	329	-----	329
Db	1825	CACGTGTTTGCCGTCCACCCCTGAGTGTCAAGCCCCAGAAATGGCTCACTGTGTTTGA	1884
Qy	329	-----	329
Db	1885	CCGGAGGCTGACCAAGTGTGTGECCTGTGCCACTATTAAGSACCCTCCCTTCTGCGTGGCC	1944
Qy	329	-----	329
Db	1945	CGCTGCCCCAGCGGTGTGAAAAGCTGACCTCTCTACATGCCCCATCTGGAAGTTCCAGAT	2004
Qy	329	-----	329
Db	2005	GAGGAGGGCGCATCCAGGCTTGCCCATCACTGCACCCACTCTGTGTGACCTGGAT	2064
Qy	329	-----	329
Db	2065	GACAAGGCTGCCCCCGCCGAGCAGAGAGGCCAGGCCCTGTGACGTCACTGTCTGTGCGGTG	2124
Qy	329	-----	329
Db	2125	GTTGGCATTTCTGTGTGTGTGTTGTTGGGGTGGTCTTTGGGATCCTCATCAAGGAGGG	2184
Qy	329	-----	329
Db	2185	CAGCAGAAGATCCGGAAGTACAGATGCGGAGACTGCTGTGAGAAACGAGACTGTGGAG	2244
Qy	329	-----	329
Db	2245	CCGCTGACACCTAGCGGAGGGAGTCCCAACAGCGGAGATGGCGATCTCTGAAGAGAG	2304
Qy	329	-----	329
Db	2305	GAGCTAGAGGAAGTGAAGTGTGTGATCTGGGCTTTTGGACAGTCTACAAGGCGATC	2364
Qy	329	-----	329
Db	2365	TGGATCCCTGATGGGGAGGAATGTGAATAATTCAGTGGCCATCAAAAGTTGTGAGGGAAAC	2424
Qy	329	-----	329
Db	2425	ACATCCCCCAAGGCCAACAAAGAAATCTTAGAGAGACATACGTGATGCTGGTGTGGGC	2484
Qy	329	-----	329
Db	2485	TGCCCATATGTCTCCCGCTTCTTGCGCATCTGCCCTGACATCCAGGTCGACGCTGGTACA	2544
Qy	329	-----	329
Db	2545	CAGCTATGGCCCTATGGGTGCTCTTAGACCATGTCCGGGAAACCGCGAGCGCTGGGC	2604
Qy	329	-----	329
Db	2605	TCCAGAGACCTGTGAACCTGTGTATGACGATTGCCAAGGAGTGAAGTACTGTGAGAGAT	2664
Qy	329	-----	329
Db	2665	GTTGGCGCTGTACACAGAGGACTTGGCCGCTCGGAAGCTGTGTCAAGAGTCCAAACAT	2724
Qy	329	-----	329
Db	2725	GTCAAAATTTACAGACTTCGGGTGCTCGGCTGCTGACATTGACGAGACAGATACCAT	2784
Qy	329	-----	329
Db	2785	GCAGATGGGGCAAGGTGCCCATCAAGTGAATGGCTGAGTCCATTTCTCCGCGGGGG	2844
Qy	329	-----	329





Db	205	GGAGCCGGAGACACCAAGTGTGACCGGACAGACATGAACCTGGGGTCCCTCCCACT	264
Oy	58	ProGluThrHisLeuAspMetLeuArgHisLeuTyGInGlyCysGlnValValGInGly	77
Db	265	CCCGAGACCCACCTGGACATGCTCCGCCAACCTTACAGAGGCTGCAGGTGTGCAGAGA	324
Oy	78	AsnLeuGluLeuThrTyLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
Db	325	AACCTGGAACTACACTTCACCTGCCCCAACAGCCAGCTGTCTCTCTGAGGATATTCAG	384
Oy	98	GluValGInGlyTyValIleuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117
Db	385	GAGGTGCAGGGCTACGTCGTCTATCGCTCACACCAAGTAGGCGAGGTCCCACTGCAGAG	444
Oy	118	LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyAlaLeuValAlaLeuAsp	137
Db	445	CTGGGATNTGGGAGGACCCACGACTCTTTGAGAGCAACTATGACCTGGCCGCTGTAC	504
Oy	138	AsnGlnAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyIleuArg	157
Db	505	AATGGAGACCCGCTGAACCATACCCCTGTCTACAGGGGCTCCCCAGAGGCTCTGGG	564
Oy	158	GluLeuGlnLeuArgSerLeuThrGlnIleLeuLysGlyValValLeuIleGlnArgAsn	177
Db	565	GAGCTGCAGCTTGCAGACCTTCACAGATCTTAAAGAGGGGTCTTGATCCAGCGGAA	624
Oy	178	ProGlnLeuCysTyArgInaAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGln	197
Db	625	CCCCAGCTCTCTACAGAGACGATTTGTGGAAAGACATCTTCCACAAAGAACACAG	684
Oy	198	LeuAlaLeuThrIleuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet	217
Db	685	CTGGCTCTCACACTGATAGAACCAACCCGCTCGGGCTGCACCCCTGTTCTCCGATG	744
Oy	218	CysLysGlySerArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThr	237
Db	745	TGTAAAGGGCTCCGCTGCTGGGGAGAGTTCTGAGGATTTGTCAAGCCTGAGCGGCACT	804
Oy	238	ValLysAlaGlyGlyCysAlaIleArgCysLysGlyProLeuProThrAspCysCysHisGlu	257
Db	805	GTCGTGCGGGTGGCTGTGCCCTGCAGAGGGGCCACTCCCACTGACTGCTGCCATGAG	864
Oy	258	GInCysAlaAlaIleGlyCysThrGlyProLysHisSerAspCysLeuAlaLacCysLeuHisPhe	277
Db	865	CAGTGTGCTCGGGCTCCACAGGGGCCCAAGCACTGTGACTGCTGGCTCCACTTC	924
Oy	278	AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyAsnThrAspThr	297
Db	925	AACCAAGTGGCATCTGTAGCTGTCACTGGCCAGCCCTGTCACTACAAACAGACAGAG	984
Oy	298	PheGluSerMetProAsnProGluGlyAlaGlyTyThrPheGlyAlaSerCysValThrAla	317
Db	985	TTTAAAGTCCAAAGCCCAATCCCGAGGGCGGTAATACATTTGGCGCCAGCTGTGTGACTGC	1044
Oy	318	CysProTyAsnTyIleLeuSerThrAspValGlySer	329
Db	1045	TGTCTCTAACAACTACTTTTCTAGGAGAGTGGATCTTGCAACCTGTGTGCCCCCTGCAC	1104
Oy	329	-----	329
Db	1105	AACCAAGAGTGAACAGACAGAGATGAAACACAGCGGTGTGAGAACTGACGAAGCCCTGT	1164
Oy	329	-----	329
Db	1165	GCCCGAGTGTCTATGTCTTGGGCATGAGACACTTGCAGAGGTGAGGGCAGTTACCACT	1224
Oy	329	-----	329
Db	1225	GCCAAATATCCAGAGATTGTGGGTGCAAAATAATCTTTTGGAGACCTGTGCATTTCTGGCG	1284
Oy	329	-----	329

Db	1285	GAGAGCTTTATATGGGGAGCCAGGCTCCAAACTGCCCCGCTCCAGCCAGAGCAATCTCCAA	1344
QY	329	-----	329
Db	1345	GTGTTGAGACTGTGAAAGAGATCACAGGTAACTTATACATCTCAGATGGCCGAGACG	1404
QY	329	-----	329
Db	1405	CTGCCTGACCTCAGCGTCTTCCAGAACTGCAAGTAATCCGGGGAGAAATTCTGCACAA	1464
QY	329	-----	329
Db	1465	GGCCCTACTGCTGACCCCTGCAAGGCGTGGGCATGAGCTGGCTGGGCTGCCTCACTG	1524
QY	329	-----	329
Db	1525	AGGGAACGGGCAAGTGGACGTCGGCCCTCATACACATAAACACCACCTCGCTGTGTGCAC	1584
QY	329	-----	329
Db	1585	ACGGTGCCCTGGGACACAGCTTTTCGGAACCGCACACAAAGCTCTGTCCACACTGCCAAC	1644
QY	329	-----	329
Db	1645	CGGCCAGAGACGAGTGTGTGGCGAGGGCTGGCCCTGCCACAGCTGTGCGCCCGAGGG	1704
QY	329	-----	329
Db	1705	CACGTGTGGGSTCCAGGGGCCACCACAGTGTGTCAACTGCAGCAGTTCTTCGGGGCCAG	1764
QY	329	-----	329
Db	1765	GAGTGCCTGAGAGAAATGCGAGTACTGACAGGGCTCCGAGGAGATATGTAAATCCAGG	1824
QY	329	-----	329
Db	1825	CACGTGTTGCCGTGCACACCTGAGTGCAGGCCAGAACATGCTCAGTCACTGTTTGGGA	1884
QY	329	-----	329
Db	1885	CCGGAGGCTGACCAAGTGTGTGGGCTGTGGCCACTATAAGAGACCCCTCCCTGTGCGTGGCC	1944
QY	329	-----	329
Db	1945	CGTGCCCAAGCGGTGTGAACCTGACCTCTCTACATGCCATGGAAGTTTCCAGAT	2004
QY	329	-----	329
Db	2005	GAGGAGGGCGCATGCAGCCTTGCCCCCATCACTGCACCCACTCTGTGTGACTGGAT	2064
QY	329	-----	329
Db	2065	GACAAAGGCTGCCCGCCGAGACAGAGAGCCAGCCCTCTACGTCCATGCTCTGCGGTG	2124
QY	329	-----	329
Db	2125	GTTGGCATTCGTGCTGTGTGTCTTGGGGTGTCTTTGGGATCCTCATCAAGCGACGG	2184
QY	329	-----	329
Db	2185	CAGCAGAAGATCCGGAAGTACACAGATGCGAGACTGCTGCAGAAACGAGACTGTGGAG	2244
QY	329	-----	329
Db	2245	CCGCTGACACCTAGCGGAGCGATGCCCAACAGGCGCAATCGCGATCTCTGAAGAGAG	2304
QY	329	-----	329
Db	2305	GAGCTGAGAGAGTGAAGGTCTTGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATC	2364
QY	329	-----	329
Db	2365	TGAGTCCCTGATGGGAGAAATGTGAATAATTCAGTGGCCATCAAGAGTGTGAGGAAAC	2424

QY 329 ----- 329  
 Db 2425 ACATCCCCAAGCAACAAAGAAATCTTAGACGAGCATACGTGATGCTGTGTGGGC 2484  
 QY 329 ----- 329  
 Db 2485 TCCCCAATATGTCTCCCGCTTCTGGGCACTGCTGACATCCAGGTGACGTGTGACA 2544  
 QY 329 ----- 329  
 Db 2545 CAGCTTATGCCCTATGGGTGCTCTTAGACCATGTCCGGAAAAACCGGGAGCGCTGGGC 2604  
 QY 329 ----- 329  
 Db 2605 TCCCAAGACCTGCTGAACCTGGTGTATGCAGATTCGAAGGGATAGCTACCTGAGAGAT 2664  
 QY 329 ----- 329  
 Db 2665 GTGGGGCTCGTACACAGGAGCTTGGCCGCTCGGAACGCTGCTGCAAGAGTCCCAACCAT 2724  
 QY 329 ----- 329  
 Db 2725 GTCAAAATATACAGACTTGGGCTGTGGCTGCTGGACATTGAGAGACAGATACCAT 2784  
 QY 329 ----- 329  
 Db 2785 GCAGATGGGGCAAGGTGCCCATCAAGTGAATGGCGCTGAGTCCATTCGCCGGCGG 2844  
 QY 329 ----- 329  
 Db 2845 TTCAACCAACAGAGATGATGTGTGAGATTATGCTGACTGTGTGGAGCTGATCACTTT 2904  
 QY 329 ----- 329  
 Db 2905 GGGCCAAACCTTAGATGGATGGATCCCAAGCCGGGAGATCCCTGACTGTGTGAAAAGGGG 2964  
 QY 329 ----- 329  
 Db 2965 GAGCGGCTGCCCAAGCCCCCATCTGCACCATATGATGTCTACATGATGTCATAATGT 3024  
 QY 329 ----- 329  
 Db 3025 TGGATGATTGACTGTGAATGTGCGCCCAAGATTCGGGAGTTGGTGTGAATTCCTCCGC 3084  
 QY 329 ----- 329  
 Db 3085 ATGGCCAGGAGCCCCAGCGCTTGTGTGTCATCCAGATGAGACTTGGGCCACGCACT 3144  
 QY 329 ----- 329  
 Db 3145 CCTTGGACAGACCTTCTACCGCTCTGCTGAGAGCATGATGGGGAGCTGGTG 3204  
 QY 330 ----- G1y 330  
 Db 3205 GATGCTGAGAGATCTGTGTGATCCCAAGAGGCTTCTTCTGTCCAGACCTGCCCGGGC 3264  
 QY 331 A1a1g1y1m2v1a1h1s1a1r1g1a1r1s1e1r1s1e1r1h1a1r1s1e1r1g1y1a1s1p 350  
 Db 3265 GCTGGGGGCATGATCCACACAGGACCGCATCTCATCTACAGAGATGGGGGTGGGAC 3324  
 QY 351 LeuTh1rLeuG1yLeuG1uP1rO1s1e1r1G1uG1u1a1P1rO1a1r1S1e1r1P1rO1e1u1a1P1rO1s1e1r 370  
 Db 3325 CTGACACTAGGGCTGAGACCTCTGAAGAGAGGCCCCCAAGCTCTCCACTGGCACCTCC 3384  
 QY 371 G1uG1y1a1a1g1y1s1e1r1a1p1a1l1a1P1rO1a1r1S1e1r1h1a1r1s1e1r1g1y1a1s1p1l1e1u 390  
 Db 3385 GAAGGGCTGGCTCGATGATATTTGATGTGACTTGGAAATGGGGGAGCCCAAGGGGCTG 3444  
 QY 391 G1u1s1e1r1e1u1P1rO1h1r1a1s1p1rO1e1u1a1r1g1y1s1e1r1G1u1a1s1p1rO1h1r1V1a1 410  
 Db 3445 CAAGGCTCCCAACATGACCCCAAGCCCTCTACAGCGGTACAGTGAAGACCCCAAGTA 3504

QY 411 ProleuP1rO1s1e1r1G1uTh1r1a1s1p1y1a1a1P1rO1e1uTh1r1C1y1s1e1r1P1rO1G1u 430  
 Db 3505 CCCCCCTGCTGTGAGACTGATGAGCTGACCTTGTCCCTGACCTGACCTGACCTGACCTGAA 3564  
 QY 431 Tyr1a1s1a1nG1uP1rO1a1r1S1e1r1a1r1g1P1rO1G1uP1rO1s1e1r1P1rO1a1r1S1e1r1P1rO1e1uP1rO 450  
 Db 3565 TATGTGAACCAAGCAGATGTGGGCCCAAGCCCTTGTCCCGCAAGAGGGCTGTGCT 3624  
 QY 451 A1a1a1r1g1P1rO1a1r1S1e1r1h1a1r1s1e1r1G1u1a1r1g1a1l1y1s1r1h1e1u1s1e1r1P1rO1G1y1a1s1n 470  
 Db 3625 GCTGCCGACCTGCTGTGGTCCCACTGTGAAAGGGCCAAAGACTCTCTCCCAAGGAAAGAT 3684  
 QY 471 G1y1a1v1a1l1y1a1s1p1a1l1a1P1h1e1g1y1a1a1v1a1G1u1a1s1p1rO1G1uTh1r1e1uTh1r 490  
 Db 3685 GGGGTGCAAAAGACGTTTGGCTTGGGGGTCCCGGAGAACCCCACTACTTGCA 3744  
 QY 491 ProG1uG1y1a1a1a1P1rO1G1uP1rO1h1a1r1S1e1r1P1rO1a1r1S1e1r1P1rO1a1P1h1e1s1P1rO1a1P1h1e1s1P 510  
 Db 3745 CCCCAGGAGAGCTGCCCCCTGACCCCACTCTCTGCTTCAAGCCAGCTTGAC 3804  
 QY 511 Asn1e1uTh1r1Y1r1T1r1a1s1p1G1u1a1s1p1rO1G1u1a1r1g1a1l1a1P1rO1s1e1r1Th1r1P1h1e1y1s 530  
 Db 3805 AACCTATATCTGAGACCAAGACCAACAGAGGGGGGCTCCACCAAGCACTTCAA 3864  
 QY 531 G1yTh1rP1rO1h1r1a1a1G1u1a1s1p1rO1G1uTh1r1e1uG1y1e1u1a1s1p1a1l1a1a1P1rO1 550  
 Db 3865 GGGACACTTACGCAAGAACCCAGAGTACCTGGGTGTGACGTGCCA-----GTGTGA 3918  
 QY 551 A1a1r1s1e1r1P1rO1s1e1r1P1rO1 556  
 Db 3919 ACCAGAAAGCCAGTCCG 3936  
 RESULT 10  
 ID ABR83918 standard; cDNA; 4530 BP.  
 AC ABR83918;  
 DT 14-AUG-2002 (first entry)  
 DE Human cDNA differentially expressed in granulocytic cells #489.  
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
 KW viral infection; parasitic infection; protozoal infection;  
 KW fungal infection; sterile inflammatory disease; psoriasis;  
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KW adult respiratory distress syndrome; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; periodontal disease;  
 KW granulocyte activation; chronic inflammation; allergy.  
 OS Homo sapiens.  
 PN W0200228999-A2.  
 PD 11-APR-2002.  
 PF 03-OCT-2001; 2001WO-US30821.  
 PR 03-OCT-2000; 2000US-237189P.  
 PA (GENE-) GENE LOGIC INC.  
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
 DR WPI; 2002-435328/46.  
 PT Detecting granulocyte activation by detecting differential expression  
 of genes associated with granulocyte activation, which serves as  
 PT diagnostic markers that is useful for monitoring disease states and  
 drug toxicity  
 PS Claim 1; SEQ ID NO 489; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. CC Also included are modulating (M2) GA by contacting GC with an agent CC that alters the expression of at least one gene in Gs: (2) screening (M3) CC for an agent capable of modulating GCA or an inflammation (especially CC chronic) in a tissue, an allergic response in a subject, exposure of a CC subject to a pathogen or sterile inflammatory disease using the CC gene expression profile; (3) detecting (M4) an inflammation (especially CC chronic) in a tissue, an allergic response in a subject, exposure of a CC subject to a pathogen or sterile inflammatory disease, by detecting the CC level of expression in a sample of the tissue of gene(s) from Gs, where CC the level of expression of the gene is indicative of inflammation, CC (4) treating (M5) an inflammation (especially chronic) or in a tissue, CC an allergic response in a subject, exposure of a subject to a pathogen CC or sterile inflammatory disease, by contacting a tissue having CC inflammation with an agent that modulates the expression of gene(s) CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for CC modulating GA; M3 is useful for screening an agent capable of modulating CC GCA preferably in an inflammation in a tissue; M4 is useful for CC detecting an inflammation (especially chronic) in a tissue, an allergic CC response in a subject, exposure of a subject to a pathogen or sterile CC inflammatory disease (e.g., psoriasis, rheumatoid arthritis, CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal CC reperfusion injury, ARDS, adult respiratory distress syndrome, CC inflammatory bowel disease, Crohn's disease, ulcerative colitis, CC periodontal disease, also bacterial infection, viral infection, CC parasitic infection, protozoal infection, fungal infection, and M5 is CC useful for treating one of the above conditions. The present CC sequence represents a gene differentially expressed in granulocytes. CC Note: The sequence data for this patent did not form part CC of the printed specification, but was obtained in electronic CC format directly from WIPO at CC ftp://ipo.int/pub/published\_pcl\_sequences.

XX Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other:

Alignment Scores:  
 Pred. No.: 8,15e-98 Length: 4530  
 Score: 2423.00 Matches: 524  
 Percent Similarity: 41.14% Conservative: 5  
 Best Local Similarity: 40.75% Mismatches: 21  
 Query Match: 64.80% Indels: 736  
 DB: 24 Gaps: 4

US-09-821-883-2 (1-690) x ABK83918 (1-4530)

QY 2 ArgAlaAlaProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17  
 ||||| ||| :|||:|:| |||:|:|:|:|  
 Db CGGCCCTCCAGCGGGGTCCAGCGGAGCCATGCGGCCGAGCCGACAGTACCAACCATG 153  
 QY 18 PheLeuPheLeuPhePheTrpLeuAspArgSerValLeuAlaAlaLysGluLeuAlaArg 37  
 ||| ||| ||| :|||:|:| |||  
 Db GAGCTGGCGGCTTGCGCCCTGG-----GGGCTCTCTCGGCCCTTGGCCCCC 204  
 QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuAlaGluProAlaSer 57  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db GGACCCCGAGACCCCAAGTGTGACCGGCGACACATGAAGCTGCGGCTTCCCTCCAGT 264  
 QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrgInglyCysGlnValValGlnGly 77  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db CCCAGACCCACCTGGACATCTCCGCCACTTACCAAGGGCTGCGAGGTGGTGCAGGA 324  
 QY 78 AsnLeuGluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db AACCTGGAACCTACCTACCTGCCACCAATGCCAGCTGCTCTTCCTGACAGATATCCAG 384  
 QY 98 GluValGlnGlyTyrrValLeuIleAlaHisAsnGlnValAlaGlnValProLeuGlnArg 117  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db GAGGTGAGGGGTACGTGCTCATGCTCACAAACAGTGAAGGAGGAGGTCCCTCCAGAGG 444

QY 118 LeuArgIleValAlaArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValIleuAsp 137  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db CTGCGGATTTGGCAGGACGCCACCACTTTTGGAGACAACTATGCCCTGGCCCTCTAGAC 504  
 QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db AATGAGACCCCGTGAACAATATACCAACCCCTGTACAGGGGGCTCCCGAGAGGCTCGGG 564  
 QY 158 GluLeuGlnLeuAspSerLeuThrGluIleLeuLysGlyValIleuIleGlnArgAsn 177  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db GAGCTGACGCTCCAAACCTCCACAGAGATCTTGAAAGAGGGGTCTTGATCCACCGAAC 624  
 QY 178 ProGlnLeuGlyTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAspAsnGln 197  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db CCCAGCTCTCTACCGAGACACATTTTGGAGAGACATCTTCCACAGACACACAG 684  
 QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db CTGCTCTACACATGATAGACACCAACCGCTCTCGGGCTCCACCCCTGTCTCCGATG 744  
 QY 218 CysLysGlySerArgCysTyrrPglYgluSerSerLysArgCysGlnSerLeuThrArgThr 237  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db TGTAAAGGGCTCCCTCTCTGGGAGAGATTCTGAGATTTCTAGACCTGACCGCACT 804  
 QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu 257  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db GTCTGTCCGGTGGCTGTGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 864  
 QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db CAGTGTGCTGCGGCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 924  
 QY 278 AsnHisSerGlyIleCysGluLeuHisCysArgProAlaLeuValIleTrpTyrrAsnThrAspThr 297  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db AACCAAGTGCATCTGTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 984  
 QY 298 PheGluSerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValIleAla 317  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db TTTGAGTCCATGCCCATTCGCCAGAGGCGGATATTCATTTCGGCGCAGCTGTCTACATGCC 1044  
 QY 318 CysProTyrrAsnTyrrLeuSerThrAspValGlySer----- 329  
 ||||| ||||| ||||| :|||:|:| |||||  
 Db TGTCCCTTAACAACCTACCTTTTACAGGAGGTGGATCTGCACCTGTCGTCCTGCGAC 1104  
 QY 329 ----- 329  
 Db AACCAAGAGGTGACAGACAGAGATGAACACAGCGGTGTGAGAACTGCAGTACGCCCTGT 1164  
 QY 329 ----- 329  
 Db GCCCGAGTGTCTATGTGCTGTGGCATGAGACACTTGGAGAGAGTGAAGGCAATTACCACT 1224  
 QY 329 ----- 329  
 Db GCCAATATCCAGAGATTGTGCTGCTGCAGAGAAGATCTTTGGAGCCTGGCATTTCTGCGC 1284  
 QY 329 ----- 329  
 Db GAGAGCTTTGTGGGGAGCCAGCCTCCAAACAGCTGCCCCGCTCCAGACAGACAGCTCCAA 1344  
 QY 329 ----- 329  
 Db GTGTTTGAAGACTTGGAAGAGATCAGAGGTACCTATATCATCTCAGCATGGCGGAGACAG 1404  
 QY 329 ----- 329  
 Db CTGCTGACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGAGAGAAATTCTGCACAA 1464  
 QY 329 ----- 329  
 Db GCGGCTACTGCTGACCCCTGCAAGGCTGGGATCTAGCTGAGCTGGGGCTGCGCTACTG 1524

329 ----- 329  
Db 1525 AGGAACTGGGAGTGTGACTGGCCCTCATCCACATACACCCACTCTGCTGTGAC 1584  
329 ----- 329  
Db 1585 ACGGTGCTGGGACACAGCTCTTTGGAACCCGACCAAGCTCTGCTCACTGCGCAAC 1644  
329 ----- 329  
Db 1645 CGGCAAGAGCAGAGTGTGTGGGAGGGGCTGGCCCTGCCACGAGTGTGCGCCGAGG 1704  
329 ----- 329  
Db 1705 CACTGTGGGGTCCAGGGGCCACCACTGTGTCACTGACGAGTCTCTTGGGGCCAG 1764  
329 ----- 329  
Db 1765 GAGTGGGTGAGGAATGCCAGTACTGACAGGGGCTCCCGAGGAGTATGGAATGCCAG 1824  
329 ----- 329  
Db 1825 CACTGTTTGGCGTGCACCCCTGAGTGTACAGCCCGAAGATGGCTCAGTACCTGTTTGA 1884  
329 ----- 329  
Db 1885 CCGGAGGCTGACCAAGTGTGTGGCTGTGCCACTATAGAGACCTCCCTTCTGTGCGCC 1944  
329 ----- 329  
Db 1945 CGGTGCCCGACGGGTGTGAACCTGACCTCTCTACATGCCCATGTGGAAGTTTCCAGAT 2004  
329 ----- 329  
Db 2005 GAGAGGGCGCATGACAGCTTGCCCATCACTGACCCACTCTGTGTGACCTGGAT 2064  
329 ----- 329  
Db 2065 GACAAAGGCTGCCCGCCGAGAGAGAGAGAGCCCTGTGACATCATGTCTGTGCGGTG 2124  
329 ----- 329  
Db 2125 GTTGGCATTTGCTGTGTGTGTGTGTGGGGGTGCTTTTGGGATCCTCATCAAGCGACG 2184  
329 ----- 329  
Db 2185 CACGAGAAATCCGGAAGTACAGATGCGGAGACTGCTGAGAAACGAGAGTGTGAG 2244  
329 ----- 329  
Db 2245 CCGCTGACACCTAGCGGAGCGATGCCAACGCGGATCGGAAAGAGAGC 2304  
329 ----- 329  
Db 2305 GAGCTGAGAAAGTGAAGTGTGTGATCTGTGGGCTTTTGGACAGTCTCAAGGGGATC 2364  
329 ----- 329  
Db 2365 TGGATCCTGATGSGGAGAAATGTGAATAATTCAGTGGCCATCAAAAGTGTGAGGAAAC 2424  
329 ----- 329  
Db 2425 ACATCCCCCAAGCAAAAGAAATCTTAGACGAAGCATAGTGTGTGTGGGC 2484  
329 ----- 329  
Db 2485 TCCCATATGTCTCCGCGCTTGTGGGATCTGCTGATCAGATCCAGGTTGAGCTGTGACA 2544  
329 ----- 329  
Db 2545 CAGCTTATGCTATGTGCTGTCTTAGACATGTCCGGGAAACGCGGAGCGCTGTGGG 2604  
329 ----- 329

Db 2605 TCCAGGACCTGTGAACTGGTGTATGACAGATTGCCAAGGGATGACTACTGAGAT 2664  
329 ----- 329  
Db 2665 GTGGGCTCTTACACAGGACTTGGCCGCTCGAAGCTGCTGCTCAAGACTCCACCAT 2724  
329 ----- 329  
Db 2725 GTCAAAATTAAGACTTCCGGGCTGGCTCGGCTGTGACATTGACAGACAGATACAT 2784  
329 ----- 329  
Db 2785 GCAGATGGGGCAAGTGGCCATCAAGTGAATGCGCTGAGTCCATTCTCCGCGCGG 2844  
329 ----- 329  
Db 2845 TTCACCCACAGAGTGTGTGTGAGTATGTGTGACTGTGTGGAGCTGATGACTTTT 2904  
329 ----- 329  
Db 2905 GGGCCAAACCTTACGATGGATCCAGCCGGGAGATCCCTGACTGTGAAAAGGG 2964  
329 ----- 329  
Db 2965 GAGCGCTGCCAGCCGCCATTCGACATTGATGTATCATGATCATGTCAAAATGT 3024  
329 ----- 329  
Db 3025 TGGATGATTACTCTGAATGTGGCCCAAGATTCGGGAGTTGTGTGAATTTCCCGC 3084  
329 ----- 329  
Db 3085 ATGGCCAGGAGCCCGAGCGCTTTGTGTATCCAGATGAGACTTGGCCAGCCAGT 3144  
329 ----- 329  
Db 3145 CCCTTGACAGACCTTTACCCGCTCAGCTGTGAGGACATGATGGGGACTGGTG 3204  
330 ----- Gly 330  
Db 3205 GATGCTGAGAGTATCTGTGATCCACAGAGGCTTCTTGTCCAGACCTGCCCGCGG 3264  
331 ----- 331  
Db 331 AlAGlyGlymeValHisHisArgSerSerThrArgSerGlyGlyLeu 3350  
3265 GCTGGGGCATGTGTCACACAGGACCGACGCTCATCTACAGAGTGGCGGGGAGC 3324  
351 LeuThrLeuGlyLeuGlyProSerGlyGlyLeuAlaProArgSerProLeuAlaProSer 370  
3325 CTGACCTAGAGGCTGAGCCCTCTGAGAGAGAGGCCCCAGGTCTCCACTGGCACCTCC 3384  
371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyAlaAlaGlyLeu 390  
3385 GAAGGGCTGGCTGCGCATGATTTGATGTGACTGGAGATGGGGGACGCAAGGGGCTG 3444  
391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGlyAspProThrVal 410  
3445 CAAAGCTCCCAACATGAGCCCAAGCCCTCTACAGGGGACAGTGGAGCCCAACGTA 3504  
411 ProLeuProSerGlyProAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 430  
3505 CCCCTGCCCTGTGAGACTGATGGCTTACGCTTGGCCCTGACCTGACGCCCCAGCTGAA 3564  
431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyGlyProLeuPro 450  
3565 TATGTGAACAGCAGATGTTGGCCCCAGCCCTTCCGCCGAGAGGGGCGCTCTCCT 3624  
451 AlaAlaArgProAlaGlyAlaThrLeuGlyArgAlaAlaGlyThrLeuSerProGlyLeuAsn 470  
3625 GCTGCCGACCTGTGTGCTGCTGGAAGGGCCAAAGACTCTCTCCAGGGAAGAAAT 3684  
471 GlyValValLeuAspValPheAlaPheGlyGlyAlaValAlaGlnAsnProGlyTyrLeuThr 490  
|||||

Db 3685 GGGGCTGTCMAAGACGTTTTCCTTGGGGGCTGCCGTGGAGACCCCGAGTACTTGACA 3744  
 Qy 491 ProglnglglYAlaAlaProglInProHsProProAlaPheserProAlaPhasp 510  
 Db 3745 CCCCAGGAGAGCTGCCCCCTTACGCCCTCTCTGCTTACGCCCTTCCGAC 3804  
 Qy 511 AsnLeuTYTTPRAspGlnAspProProgluArglYAlaProPserThrhPhuys 530  
 Db 3805 AACCTCTATTACTGGACAGACAGACAGAGGGGGGCTCCACCCAGCACTTCAAA 3864  
 Qy 531 GlyThrProThAlaGlnAsnProgluTYrlLeuAlaSpvalProAlaAlaPro 550  
 Db 3865 GGGACACCTACGCGAGAGAACCGAGTACCTGGCTGTGAGAGTCCA-----GTGTGA 3918  
 Qy 551 AlaArgSerProSerPro 556  
 Db 3919 ACCAGAGGCCCAAGTCCG 3936  
 RESULT 11  
 ABK14057  
 ID ABK14057 standard; DNA; 9274 BP.  
 AC ABK14057;  
 DX 23-APR-2002 (first entry)  
 XX  
 DE Human HER2 (ErbB2) transgene plasmid construct DNA.  
 XX  
 KW Human; HER2; ErbB; epidermal growth factor receptor; gene; ds;  
 KW anti-ErbB antibody-maytansinoid conjugate; cancer; tumour; breast; ovary;  
 KW stomach; endometrium; salivary gland; lung; kidney; colon; colorectum;  
 KW thyroid; pancreas; prostate; bladder; neuronal disorder;  
 KW glioma disorder; astrocytic disorder; hypothalamic disorder;  
 KW glandular disorder; macropneal disorder; epithelial disorder;  
 KW stroma disorder; blastocoele disorder; inflammatory disorder;  
 KW angogenic disorder; immunological disorder.  
 KW  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN US2002001587-A1.  
 XX  
 PD 03-JAN-2002.  
 PF 16-MAR-2001; 2001US-081123.  
 PR 16-MAR-2000; 2000US-189844P.  
 PR 05-OCT-2000; 2000US-238527P.  
 XX  
 PA (ERIC/) ERICKSON S.  
 PA (SCHW/) SCHWALL R.  
 PA (SLIW/) SLIWOMSKI M.  
 XX  
 PI Erickson S, Schwall R, Sliwowski M;  
 DR WPI: 2002-163686/21.  
 DX  
 PT Treating tumour characterised by overexpression of epidermal growth  
 PT factor receptor, ErbB or cancer in mammal, comprises administering  
 PT anti-ErbB antibody-maytansinoid conjugate to the mammal  
 XX  
 PS Example 3; Fig 7; 93pp; English.  
 CC The invention relates to treating a tumour in a mammal, where the tumour  
 CC is characterised by the overexpression of an epidermal growth factor  
 CC receptor (ErbB) and does not respond or responds poorly, to treatment  
 CC with an anti-ErbB antibody, comprising administering to the mammal an  
 CC anti-ErbB antibody-maytansinoid conjugate. The method is useful for  
 CC treating cancer or tumours of the breast, ovary, stomach, endometrium,  
 CC salivary gland, lung, kidney, colon, colorectum, thyroid, pancreas,  
 CC prostate and bladder, preferably breast cancer. The breast cancer is a  
 CC metastatic breast cancer or an aggressive form of metastatic breast  
 CC cancer which overexpresses ErbB2. The method is also useful for treating

CC neuronal, glial, astrocytic, hypothalamic, glandular, macropneal,  
 CC epithelial, stromal, blastocoele, inflammatory, angogenic (ErbB2)  
 CC immunological disorders. This sequence represents a HER2 (ErbB2)  
 CC transgene plasmid construct used to direct expression of native human  
 CC HER2 in the mammary gland of a transgenic mouse.  
 XX  
 SQ Sequence 9274 BP; 2152 A; 2538 C; 2489 G; 2095 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1.76e-97 Length: 9274  
 Score: 2422.00 Matches: 528  
 Percent Similarity: 41.12% Conservative: 9  
 Best Local Similarity: 40.43% Mismatches: 31  
 Query Match: 64.78% Indels: 738  
 DB: Gaps: 6  
 US-09-821-883-2 (1-690) x ABK14057 (1-9274)  
 Qy 2 ArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 21  
 Db 1695 CGCCCCGGCCCTCCAGCGGGGTCCAGCGGAGCCATGAGACTGGCG-----GCC 1745  
 Qy 22 LeuphePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSer 41  
 Db 1746 TGTGCTCCCTCG-----GGGCTCTCTCTCGCCCTCTTCCCGCCGAGCCCGAGC 1796  
 Qy 42 ThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHis 61  
 Db 1797 ACCCAAGTGTGACCGGCGCACACATGAAAGTGGCGTCCCTCCGACATCCAGACCCAC 1856  
 Qy 62 LeuAspMetLeuArgHisLeuTYrlGlnGlyCysGlnValAlaGlnGlyAsnLeuGluLeu 81  
 Db 1857 CTGACATGCTCTCGCACCTTACACAGGGCTCCAGAGTGGTCCAGAGAACTGGAACTC 1916  
 Qy 82 ThrTYrlLeuProThAsnAlaSerLeuSerPheLeuGlnAspIleGlnValAlaGlnGly 101  
 Db 1917 ACCCTACCTGCCACCAATGCCAGCTGTCTTCCTGACAGATATCCAGAGGTGACAGGC 1976  
 Qy 102 TYrlValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgGluArgIleVal 121  
 Db 1977 TACGTGCTCATGCTGCACACCAAGTAGAGGCTCCCACTGCACAGCTGGGATTGG 2036  
 Qy 122 ArgGlyThrGlnLeuPheGlnAspAsnTYrlAlaLeuAlaValLeuAspAsnGlyAspPro 141  
 Db 2037 CGAGGACCCACAGCTCTTGTAGACAACTATGCCCTGGCGCTGTAGAAATGAGACCCG 2096  
 Qy 142 LeuAsnAsnThrThProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu 161  
 Db 2097 CTGAACAAATACCAACCCCTGTCCACAGGGGCTCCCGAGAGGCTCGGAGCTGACGTT 2156  
 Qy 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeuGly 181  
 Db 2157 CGAAGCTTCACAGAGATCTTGAAGAGGGGCTTGTATCCAGAGGAAACCCCACTCTGC 2216  
 Qy 182 TYrlGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201  
 Db 2217 TACCAAGACAGATTTTGGAGAGACACTTCCAAAGAAACACAGCTGGCTGCACA 2276  
 Qy 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221  
 Db 2277 CTGATAGACACCAACCGCTCTGGGCTGCCACCCCTGTTCCTCCGATGTGTAAGGGCTTC 2336  
 Qy 222 ArgCysTrpGlyGlySerSerGlnAspCysGlnSerLeuThrArgThrValCysAlaGly 241  
 Db 2337 CCTGCTGGGAGAGAGATTCTAGAGATTGTCAAGCTTGAACCGCACTGTCTGTGCGGT 2396  
 Qy 242 GlyCysAlaArgCysLysGlyProLeuProThAspCysCysHisGluGlyCysAlaAla 261  
 Db 2397 GGCTGTGGCGGCTGCAAGGGGCACTGCCCACTGACTGTGSCAAGAGAGTGTGGCC 2456  
 Qy 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281  
 Db 2457 GGCTGCACGGGCCCAAGACACTGTGACTGCCCTGGCTGCTCCACTTCAACACAGTGGC 2516

QY 282 ILeCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMet 301  
Db 2517 ATCGTGTAGCTGCACCTCCGACCCCTGCTGACCTACACACAGACAGCTTTGATCCATG 2576  
QY 302 ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn 321  
Db 2577 CCCAATCCCGAGGCGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTCCCTACAC 2636  
QY 322 TyrLeuSerThrAspValGlySer 329  
Db 2637 TACCTTTCTACGAGCTGGAGTCTGCACCTGCTGCCCCCTGCACACACAAAGAGTG 2686  
QY 329 329  
Db 2697 ACACAGAGATGTGACACACAGCGGTGTGAGAAATGTGACAAAGCCCTGTGCCGAGTGTGC 2756  
QY 329 329  
Db 2757 TATGCTGTGGGAGAGAGACCTTGCAGAGGTGAGGCGAGTTACAGTGCATATATCCAG 2816  
QY 329 329  
Db 2817 GAGTTTCTGCTGCTCAGAAAGATCTTTGGAGCTGCGCATTTCTGCCGAGAGCTTTGAT 2876  
QY 329 329  
Db 2877 GGGAGCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGAGCTCAAGTGTGAGACT 2936  
QY 329 329  
Db 2937 CTGAGAGATACAGGTACCTATACATCTCAGCATGGCCGAGACGCTGCTGACCTC 2996  
QY 329 329  
Db 2997 AGCCTTCCAGAACTGCAAGTATCCGGGAGCAATTCTGCACATGGCGCTACTCG 3056  
QY 329 329  
Db 3057 CTGACCTGCAGAGGCTGGGATCAGCTGGCTGGGGCTGGCTCAGAGGAACTGGGC 3116  
QY 329 329  
Db 3117 AGTGGACTGGCCCTCATCCACCATACACCCACTGCTTCTGTCACAGCGTCCCTGG 3176  
QY 329 329  
Db 3177 GACCAGCTTTTGGAAACCGCACCAAGCTGTCTCACACTGCCAACCGGCCAGAGGAC 3236  
QY 329 329  
Db 3237 GAGTGTGTGGGCGAGGGCTGGCTGGCCACAGCTGTGCGCCGAGGGGCACTGCTGGGT 3296  
QY 329 329  
Db 3297 CCAAGGCCCAACCACTGTGTCACTGCAGCCAGTTCTTGGGGGCGAGAGTGGTGAG 3356  
QY 329 329  
Db 3357 GAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGTTGCCG 3416  
QY 329 329  
Db 3417 TGCCACCCTAGTGTCAAGCCCAAGAAATGCTCAGTGAACCTTTTGGACGGAGAGTGAC 3476  
QY 329 329  
Db 3477 CAGTGTGTGGCTGTGCCCACTATAGAGACCTTCCCTTGTGCTGGCCGCTGCCAGC 3536  
QY 329 329  
Db 3537 GGTGTGAACCTGACCTCTCTCATATGCCCATCTGGAAGTTTCCAGATGAGAGGGGCA 3596

QY 329 329  
Db 3597 TGCCAGCCTTGCCCCCATCACTGCACCCACTCTGTGTGAGACTGTGATGACAAAGGCTGC 3656  
QY 329 329  
Db 3657 CCCCGCAGAGAGAGAGCCAGCCCTGTGACGTCCATCTCTGTGGGTGGTGGCATTCG 3716  
QY 329 329  
Db 3717 CTGCTGTGTCTTGGGGTGTCTTTGGATCCTCATCAAGCGACGCGACAGAAATC 3776  
QY 329 329  
Db 3777 CGAAGTACAGATGCGGAGACTGCTGCAGAAAGAGAGCTGTGTGAGCCGCTGACACT 3836  
QY 329 329  
Db 3837 AGCGAGCGATGCCCAACAGGCGCAGATGCGATCTTGAAGAGAGAGCTGAGGAG 3896  
QY 329 329  
Db 3897 GTGAAGTGTGTGATCTGGCGCTTTTGGCAGACTTACAAGGGCATCTGATCCCTGAT 3956  
QY 329 329  
Db 3957 GGGAGAAATGAAATTCAGTGGCCATCAAAAGTTGAGGAAACACATCCCCCAA 4016  
QY 329 329  
Db 4017 GCCAACAAGAAATCTTAGACGAACATACGTATGCTGTGTGGCTCCCATATGTC 4076  
QY 329 329  
Db 4077 TCOCGCTTGTGGCATCTGCTCACAATCCAGGTGAGAGTGTGACAGACTTATGCC 4136  
QY 329 329  
Db 4137 TATGCTGCTCTTAGACATGTCCGGGAAACCGGAGCGCTGGCTCCAGAGACTG 4196  
QY 329 329  
Db 4197 CTGAAGTGTGTATGACGATTTGCCAAGGGATGAGTAACTGGAGATGTGGGCTGTA 4256  
QY 329 329  
Db 4257 CACAGGACTTGGCCGCTCGGAACGTGTGTCAAGATCCCAACATGTCAAAATTACA 4316  
QY 329 329  
Db 4317 GACTTGGGCTGGCTCGGCTGTGGACATTGACGAGACAGAGTACATGATGGGGC 4376  
QY 329 329  
Db 4377 AAGTGCCTCAAGTGGATGGCTGGAGTCAATTCTCCGCGCGGCTTCAACCACAG 4436  
QY 329 329  
Db 4437 AGTATGTGTGAGTATATGTGTACTGTGTGAGAGCTGATGACTTTTGGGCCCAACT 4496  
QY 329 329  
Db 4497 TACGATGGATCCAGCGCGGAGATCCGTACCTGTGAAAGGGGAGCGGCTGCC 4556  
QY 329 329  
Db 4557 CAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGATGATTTGATGAC 4616  
QY 329 329  
Db 4617 TCTGATGTGGCCAGATTCGGGAGTGTGTCTGAATTCCTCCGATGGCCAGGAG 4676  
QY 329 329

```

Db 4677 CCCAGCCTTGTGTCATCATCAGATGAGACTTGGCCAGCAGTCCCTGGACAGC 4736
Qy 329 -----
Db 4737 ACCTTCAACCGCTCAGCTGAGAGACATGATGGGGACCTGGTGAATGCTGAGAG 4796
Qy 330 -----
Db 4797 TATCGTATACCCAGCAGGCGCTTCTTCTGTCAGACCTGCCCCGGGGCTGGGCGATG 4856
Qy 335 VALHSHLSARGHLSARSSerSerSerThrArgSerGlyGlyGlyAspLeuThrLeuGly 354
Db 4857 GTCCACACAGGACACCGCAGCTCATACAGAGAGTGGCGGGGAGCTGACACTAGAG 4916
Qy 355 LeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGlyAlaGly 374
Db 4917 CTGGAGCCCTCTGAAAGAGAGGCCCCCAGGTCTCCAGTGGACCCCTCGAAGGGGCTGGC 4976
Qy 375 SerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuPro 394
Db 4977 TCCGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5036
Qy 395 ThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeuProSer 414
Db 5037 ACACATGACCCAGCCCTCTACAGCGGTACAGTGAGAGACCCACAGTACCCCTGCCCTCT 5096
Qy 415 GlnThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAsnGln 434
Db 5097 GAGACTGATGGCTACGTTGGCCCCCTGACCTGACCTGACCCCTCGAATATGTGAAACAG 5156
Qy 435 ProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAlaArgPro 454
Db 5157 CCAGATGTTGGCCCCCAGCCCTTCCGCCAGAGAGGCCCCCTGCTGCTCCCGACT 5216
Qy 455 AlaGlyAlaThrLeuGluArgAlaAlaGlyThrLeuSerProGlyLysAsnGlyValValLys 474
Db 5217 GCTGGTGCACCTCTGAAAGAGGCCCAAGACTCTCTCCAGGAGAAATGGGGTCTCAAA 5276
Qy 475 AspValPheAlaPheGlyGlyAlaValAlaGluAsnProGluTyrLeuThrProGlnGly 494
Db 5277 GACGTTTTTGGCTTTGGGGGGTCCCTGGAGAAACCCGAGTACTTACACCCAGGAGAGA 5336
Qy 495 AlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyrTyr 514
Db 5337 GCTGCCCTCAGCCCCCAGCTCTCTGCTTCCAGCCAGCTTGAGACACTCTAATTAC 5396
Qy 515 TrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThrProThr 534
Db 5397 TGGGACACAGACCCAGAGAGGGGGCTCCAGCCAGACCTTCAAGGAGACACTAGC 5456
Qy 535 AlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSerPro 554
Db 5457 GCAGAGAACCCAGACTGCTGGTGTGACGTGCCA-----GTGTGAACCGAAGAGGCA 5510
Qy 555 SerProSerThrGlnProTropGluHisValAsnAlaIleGlnGluAlaAlaArgArgLeu 574
Db 5511 AGTCCGAG-----AAGCCTGATGTGTCTCTCAG-----GACAGAGGAAGGGGCTCTG 5561
Qy 575 AsnLeuSerArgAspThr 580
Db 5562 AGCTATTCCAGAAATGAT 5579

```

RESULT 12  
ID AAA14812 standard; cDNA: 4472 BP.

AC AAA14812;

DT 08-AUG-2000 (first entry)

DE cDNA encoding the SPLICE erB-2 receptor protein.

```

KW SPLICE erB-2 receptor protein; cell transformation disorder; cancer;
KW tumor cell proliferation; tissue degeneration; arthropathy;
KW bone resorption; inflammatory disease; degenerative disorder;
KW wound healing; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 175..3942
XX FT /*tag=a
XX FT /product= "SPLICE erB-2 protein"
XX
XX WO200020579-A1.
XX
XX 13-APR-2000.
XX
XX 01-OCT-1999; 99WO-CA00912.
XX
XX 02-OCT-1998; 98US-0165192.
XX
XX (UYMC-) UNIV MCMASTER.
XX
XX Muller WJ, Siegel PM;
XX
XX WP1: 2000-303768/26.
XX
XX P-PSDB: AAY84780.
XX
XX
XX Nucleic acid encoding an erB-2 receptor protein designated SPLICE
XX erB-2. Inhibitors of the protein are useful for treatment of cancer -
XX
XX Claim 4: Fig 1: 60pp; English.
XX
XX The present sequence encodes a SPLICE erB-2 receptor protein. The
XX protein has an in-frame deletion of 16 amino acids, 2 of which are
XX conserved cysteine residues, compared to the unspliced protein. The
XX erB-2 polynucleotide is used to construct probes for detecting
XX disorders of cell transformation such as cancer. Antibodies to the
XX protein may be used to detect SPLICE erB-2 in a sample. Agents
XX (e.g. antisense oligonucleotides) which inhibit the expression of
XX SPLICE erB-2 are useful for reducing tumor cell proliferation and
XX treating cancer. Substances which stimulate SPLICE erB-2 are useful
XX for treating conditions involving damaged cells including conditions
XX in which degeneration of tissue occurs, such as arthropathy, bone
XX resorption, inflammatory diseases, degenerative disorders of the
XX central nervous system and wound healing.
XX
XX Sequence 4472 BP: 902 A; 1383 C; 1328 G; 859 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.33e-97 Length: 4472
XX Score: 2418.00 Matches: 523
XX Percent Similarity: 41.06% Conservative: 5
XX Best Local Similarity: 40.67% Mismatches: 22
XX Query Match: 64.67% Indels: 736
XX DB: 21 Gaps: 4
XX
XX US-09-821-883-2 (1-690) x AAA14812 (1-4472)
XX
XX Qy 2 ArgAlaAlaProLeuLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 17
XX Db 118 CGCGCCCTCCAGCGCGGTCCAGCGGAGCCATGGGGCCGAGCGCAGTGAACCATG 177
XX Qy 18 PheLeuPheLeuLeuPhePheThrLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
XX Db 178 GAGCTGGCGGCTTGTGCCGCTGG-----GGCTCTCTCTGCGCTTGGCCCCC 228
XX Qy 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
XX Db 229 GGAGCGCGGAGAGCACCAAGTGTGACACGCGACAGATGAAGACTCGGCTCCCTCAGT 288
XX Qy 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
XX Db 289 CCGAGACCCACCTGACATGCTCCGACACCTTACAGAGGGGTGCCAGGTGTGTCAGAGA 348

```

QY	78	AsnLeuGluLeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
Db	349	AACTCGGAACCTACACTCTGCCACCACCAATGACAGCTCTCTCTCGTAGAGATTACAG	408
QY	98	GluValGlnGlyTrpValLeuLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117
Db	409	GAGGTGCAGGGCTACGTGCTCATCGCTCACACCAAGTAGAGGAGGTCCCACTGCAGAGG	468
QY	118	LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAsp	137
Db	469	CTGGCGATTGTGCGAGGACACCCAGCTCTTTGAGAGCAACATATGCCCTGGCCGTGTACAC	528
QY	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157
Db	529	AATGGAGACCCGCTGAACCAATACCACCCCTGTACAGAGGGGCTCCCAAGAGGCTCTGGG	588
QY	158	GluLeuGlnLeuAspSerLeuThrGlnIleLeuLysGlyIValLeuIleGlnArgAsn	177
Db	589	GAGCTGCAGCTTCGAACCTCACAGAGATCTTGAAGAGAGGGGTCTTATCCAGCGGAAAC	648
QY	178	ProGlnLeuCysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln	197
Db	649	CCCCAGCTCTGCTCCAGAGACAGAGATTGTGGAAAGACATCTCCACAAGACAAACAG	708
QY	198	LeuAlaLeuThrLeuIleAspThrAsnAspSerArgAlaCysHisProCysSerProMet	217
Db	709	CTGGCTCTACACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATG	768
QY	218	CysLysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerIleuThrArgThr	237
Db	769	TGTAAAGGCTCCCGCTCGTGGGAGAGAGTTTGCAGGATTGTACAGCCTGACGGGCACT	828
QY	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisIleGlu	257
Db	829	GCTCTGTCCGCTGCTGTGCCCTGCAGAGGGGCCACTCCCACTGACTGCTGCCATTAG	888
QY	258	GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe	277
Db	889	CAGTGTGCTGCCGCTGCACGGGCCCAAGACATCTGACTGCTGGCTGCCTCCATCTC	948
QY	278	AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThr	297
Db	949	AACCAACAATGGCAATCTGTGAGCTGCACCTGCCCGCTGTCACTACAAACACAAACAG	1008
QY	298	PheGluSerMetProAsnProGluGlyArgGlyTrpPheGlyAlaSerCysValThrAla	317
Db	1009	TTTAGTGCATAGCCCAATCCCGAGGGCCGGTATACATTTCGGCCCAAGCTGTGTACTGCC	1068
QY	318	CysProGlyAsnTrpLeuSerThrAspValGlySer-----	329
Db	1069	TGTCCCTACAACTACCTTTCTACAGAGAGTGGAGTCTGCACCCCTGTCTGCCCCCTGCAC	1128
QY	329	-----	329
Db	1129	AACCAAGAGGTGACAGCAGAGATGGAACAACACCGGTGTGAGAACTGCAGCAAGCCCTGT	1188
QY	329	-----	329
Db	1189	GCCCGAGTGTCTATTGGTCTGGGCAATGAGACACTTCGAGAGAGTAGAGGCAATTACCACT	1248
QY	329	-----	329
Db	1249	GCCAAATATCCAGAGATTGTCTGGCTGCAGAAAGAAATTTGGAGAGCCTGGCAATTTCTGGCG	1308
QY	329	-----	329
Db	1309	GAGAGCTTTGATGGGACCCAGCCCTCAACAACTGCCCGCTCCAGACAGACAGCTCCAA	1368
QY	329	-----	329
Db	1369	GTGTTTAGACTCTGGAAAGAGATCACAGGTTACATATACATCTCAGCACTGGCCGGACAGC	1428

OY	329	-----	329
Db	1429	CTGCCTGACCTCAGCGCTCTTCCAGAACCTGCAGATATCCGGGAGCAATTTCACAAAT	1488
OY	329	-----	329
Db	1489	GCGCGCTTACTGCTGACCCCTGCAAGGGCTGGGATCAGGCTGGGCTGGCGCTCAGTG	15448
OY	329	-----	329
Db	1549	AGGGAACTGGGCACTGGAGCTGGCCCTCATCCACCATACAGCCACCTCTGCTTCGTGAC	1608
OY	329	-----	329
Db	1609	ACGGTGCCCTGGGACCAAGCTCTTTCCGAACCCGACCAAGCTCTGCTCCACACTGCCAAC	1668
OY	329	-----	329
Db	1669	CGGCGAGAGACAGATGTGTGGGGAGAGGGCTGGCTGGCACACGCTGTCGCGCCGAGGG	1728
OY	329	-----	329
Db	1729	CAGTGTGGGGTCCAGGGGCCACCACAGTGTGTCACTGAGCAGCAGTCTCTTGGGGGCAG	1788
OY	329	-----	329
Db	1789	GAGTCCGTGGAGAACTCCCACTACTGCAGGGGCTCCCAAGGAGATATGTGAATCCAGG	1848
OY	329	-----	329
Db	1849	CAGTGTTCGCGCTGCACACCTGAGTGTACAGCCCAAGATGGCTCAGTGACTGTGTTTGA	1908
OY	329	-----	329
Db	1909	CCGAGGCTGACACAGTGTGTGGCTGTGCCACTATAGGACCCCTCCCTTCTGCGTGGCC	1968
OY	329	-----	329
Db	1969	CGCTGCCCAAGCGGTGTGAAACCTGACCTCTCTCATATGCCATCTGGAAGTTTCCAAAT	2028
OY	329	-----	329
Db	2029	GAGAGGGCGCATGCAGACCTTGCCCCATCACTGACCCACTGCTGTGTGGACTGGAT	2088
OY	329	-----	329
Db	2089	GACAAAGGCTGCCCGCCGAGCAGAGAGCAGCCCTCTGACGTCCATCATCTTCGCGGTG	2148
OY	329	-----	329
Db	2149	GTTGGCATTTCTGCTGGTCGTGTGCTTGTGGGGTGTCTTTGGGATCCTCATCAAGGACGG	2208
OY	329	-----	329
Db	2209	CAGCAGAAGATCCGAGGTACACGATGCGGAGACTGCTGCAGGAAACGAGCTGTGGAG	2268
OY	329	-----	329
Db	2269	CCGCTGACACTAGCGGAGGATGCCCAACAGCGCGCAGATGCGGATCCTGAAGAGACG	2328
OY	329	-----	329
Db	2329	GAGCTGAGAGAGTGTAAGTGTGTGGATCTGGCGCTTTTGGCAGACGTCTACAAAGGCATC	2388
OY	329	-----	329
Db	2389	TGGATCCCTGATGGGAGAAATGTGAAAATTTCCAGTGGCCATCAAAAGTGTGAGGGAAC	2448
OY	329	-----	329
Db	2449	ACATCCCCAAAGCAACAAGAAATCTTAGAGAGAAAGCATACGTGATGGCTGTGTGGGC	2508
OY	329	-----	329



Db 2509 TCCCATATGTCCTCCGCTTCTGGGCATCTGCTGACATCCAGCGTGCAGCTGTGACA 2568  
 Oy 329 ----- 329  
 Db 2569 CAGCTTATGCTTATGCTGCTTCTAGACATGTCGGGAAACCGCGGACGCTGGGC 2628  
 Oy 329 ----- 329  
 Db 2629 TCCGAGACCTGCTGACTGCTGTATGCAATTGCCAAGGGGATGAGTACTGAGAGAT 2688  
 Oy 329 ----- 329  
 Db 2689 GTGCGCTGCTACACAGGAGACTTGGCCGCTCGAAGCGTGTGTCAGAGTCCAAACAT 2748  
 Oy 329 ----- 329  
 Db 2749 GTCAAAATTCAGACTTGGGCTGGCTGGCTGCTGAGCATTTGACAGACAGATCAAT 2808  
 Oy 329 ----- 329  
 Db 2809 GCAGATGGGGGCAAGCTGCCCATCAAGTGGATGGCGCTGGAATTCATTTCTCCCGCGG 2868  
 Oy 329 ----- 329  
 Db 2869 TTACACCAACAGATGATGTGTGAGATTATGCTGTGACTGTGTGGAGCTGATGACTTTT 2928  
 Oy 329 ----- 329  
 Db 2929 GGGGCCAAACCTTACAGATGGATGCCAGCCCGGAGATCCCTGACTGCTGGAAAAAGGG 2988  
 Oy 329 ----- 329  
 Db 2989 GAGCGCTGCCCGACGCCCATCTGCACATTGATGCTACATGATCATGCTCAAAATGT 3048  
 Oy 329 ----- 329  
 Db 3049 TGGATGATTGACTGTGAATGTGCGCAAGATTCGGGAGTGTGTGAATTTCTCCGC 3108  
 Oy 329 ----- 329  
 Db 3109 ATGCCAGGAGACCCCGACCGCTTTGTGTGATCCAGAAATGAGACTTGGGCCACCACT 3168  
 Oy 329 ----- 329  
 Db 3169 CCTTGGACAGCACTTCTACCGCTCAGCTGCTGAGAGCATGAGGGGACCTGCTG 3228  
 Oy 330 -----Gly 330  
 Db 3229 GATGCTGAGAGTATCTGTGACCCAGCAGGGCTTCTTCTGTCAGACCTGCCCGGCG 3288  
 Oy 331 AAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3348  
 Db 3289 GCTGGGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3348  
 Oy 3351 LeuThrLeuGlyLeuGlyProSerGluGluGluGluGluGluGluGluGluGluGluGlu 3408  
 Db 3349 CTGACACTAGAGGCTGAGGCTCTCTGAGAGAGAGGCGCCAGGCTCTCCACTGACCTTCC 3408  
 Oy 371 GlnGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu 390  
 Db 3409 GAAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3468  
 Oy 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410  
 Db 3469 CAAGACCTTCCACATGATACCCAGCCCTTACAGCGGTATAGTGAAGACCCACAGTA 3528  
 Oy 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 430  
 Db 3529 CCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3588  
 Oy 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 450

Db 3589 TATGTGAACACAGATGTTGCGCCAGCCCTTCCCGCCAGAGAGGCGCTCTGCT 3648  
 Oy 451 AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyAsn 470  
 Db 3649 GCTGCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3708  
 Oy 471 GlyValIValLysAspValPheAlaPheGlyGlyAlaValAlaGluAsnProGlyLeuThr 490  
 Db 3709 GGGGCTGCTCAAGAGCTTTTCCCTTGGGGGTGCCGCTGAGAACCCGAGTACTTGACA 3768  
 Oy 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510  
 Db 3769 CCCCAGGAGAGCTGCCCTCAGCCCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3828  
 Oy 511 AsnLeuTyrTyrThrAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530  
 Db 3829 AACCTCTATTACTGGAGCAGGACACACAGCGGGGGCTCCACCAAGCATCTTCAAA 3888  
 Oy 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaPro 550  
 Db 3889 GGGACACCTAGCGGCGAGACACCAAGTACCTGGGTCTGGAGCTGCCA-----GTGCA 3942  
 Oy 551 AlaArgSerProSerPro 556  
 Db 3943 ACCAGAGGCCAAGTCCC 3960  
 RESULT 13  
 AAZ31071  
 ID AAZ31071 standard; DNA; 4473 BP.  
 AC AAZ31071;  
 XX 17-JAN-2000 (first entry)  
 DT 17-JAN-2000 (first entry)  
 DE HER-2 nucleic acid sequence.  
 XX HER-2; c-neu; ErbB2; transmembrane receptor; tyrosine kinase activity;  
 KW epidermal growth factor receptor; EGFR; HER-1; cancer; breast cancer;  
 KW ovarian cancer; gastric cancer; antisense oligonucleotide; expression;  
 KW hyperproliferative disease; ss.  
 XX Homo sapiens.  
 OS W03948906-A1.  
 PN 30-SEP-1999.  
 PD 25-MAR-1999; 99WO-US06492.  
 PF 26-MAR-1998; 98US-0048804.  
 PR (ISIS-) ISIS PHARM INC.  
 PA (PENN-) PENN STATE RES FOUND.  
 XX Bennett CF, Lipton A, Wilters LM;  
 DR WPI: 1999-610749/52.  
 PT New antisense sequences used to treat hyperproliferative conditions,  
 PT especially cancer -  
 PS Examples; Page 38-39; 44pp; English.  
 XX This is the human HER-2 polynucleotide sequence. The HER-2 gene also  
 CC called c-neu and ErbB2, encodes a transmembrane receptor, with tyrosine  
 CC kinase activity. HER-2 is related to the epidermal growth factor  
 CC receptor (EGFR or HER-1). Aberrant HER-2 expression is present in a wide  
 CC number of cancers, especially breast, ovarian and gastric cancers. This  
 CC sequence is used in the invention to design 12-25 nucleotide  
 CC oligonucleotides that decrease the expression of human HER-2. The  
 CC oligonucleotides of the invention (AAZ31067-231070) can also be used for  
 CC modulating the expression of human epidermal growth factor receptor. The  
 CC oligonucleotides are used to treat diseases or conditions associated with





```

FT      /*tag- a
XX      /product- "Human Her-2 protein"
XX      MO200222636-A1.
XX      21-MAR-2002.
XX      12-SEP-2001; 2001MO-US28572.
XX      15-SEP-2000; 2000US-0663834.
XX      (ISIS-) ISIS PHARM INC.
XX      Bennett CF, Cowser LM;
XX      WPI; 2002-471192/50.
XX      P-PSDB; AAE24067.
XX      Novel antisense oligonucleotide which modulates the expression of Human
XX      Epidermal Growth Factor receptor, Her2, is useful for treating tumors
XX      inflammation or to prevent infection in humans -
XX      Example 13; Page 94-101; 116pp; English.
XX      The invention relates to antisense compounds targetted to a nucleic
XX      acid molecule encoding Her2 (human Epidermal Growth Factor receptor 2)
XX      that specifically hybridizes with and inhibits the expression of Her2.
XX      Antisense compounds of the invention are used for treating diseases or
XX      conditions associated with Her2 such as hyperproliferative disorders
XX      e.g. lung, breast, gastric, oesophageal, colon, bladder, salivary,
XX      neural or cardiac cancer. They are also useful prophylactically e.g.
XX      to prevent or delay infection, inflammation and tumour formation. The
XX      invention is also used in gene therapy. The present sequence is human
XX      Her-2 DNA.
XX      Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other:
XX
XX      Alignment Scores:
XX      Pred. No.: 1.33e-97 Length: 4473
XX      Score: 2418.00 Matches: 523
XX      Percent Similarity: 41.06% Conservative: 5
XX      Best Local Similarity: 40.67% Mismatches: 22
XX      Query Match: 64.67% Indels: 736
XX      DB: 24 Gaps: 4
XX
XX      US-09-821-883-2 (1-690) x AAD38904 (1-4473)
XX
XX      2 ArgAlaIaIaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
XX      118 CGGCGCTCCAGCGGGGTCCAGCGGAGCCATGGGCGCGGAGCCGAGTGAACCATG 177
XX      18 PheLeuPheLeuPhePhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
XX      178 GACCTGGCGGCTGTGCTGCGCTG-----GGGCTCTCTCCGCGGCTCTTGGCCCCC 228
XX      38 GlyAlaIaIaSerThrGlnValCysThrGlyThrAspMetLysLeuAlaGluProAlaSer 57
XX      229 GGAAGCCCGGAGCAACCAAGGTGTGACCGGACAGACATGAGCGGCTCCCTGGCCAGT 288
XX      58 ProGluThrIleuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
XX      289 CCGGAGACCACTGGACATGCTCCGACCTTACCAAGGGCTGCCAGGGTGGTCAAGGA 348
XX      78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
XX      349 AACCTGGAACCTACACTACCTGCGCCACCAATGCCAGCTGCTTCCTGCGAGATATCAG 408
XX      98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
XX      409 GAGGTGAGGGGTACAGTGCATCGCTCACACCAAGTGAAGCGAGGTCCCACTGGCAAGG 468
XX      118 LeuArgIleValAlaArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137

```

```

Db      469 CTGGGAGTTGTGGAGGACCCACCTCTTGGAGGACAACTATGCCCTGGCCCTGTAC 528
XX      138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
XX      529 AATGAGAGCCCGGTGAACAATATACACCCCTGTACAGGGGGCTCCAGAGGGCTCGGG 588
XX      158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsn 177
XX      589 GAGCTGACGCTTCGAAACCTCAGACAGATCTTGAAGAGAGGCGTCTGATCCACCGAAC 648
XX      178 ProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheNHisLysAsnGln 197
XX      649 CCGGAGCTGTGCTACAGAGACAGATTTGTGAAGAGACATCTTCCAAAGAACACAG 708
XX      198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
XX      709 CTGGCTCTACACATGATAGACACCAACCGCTCTGGGGCTGGCCACCTGTCTCGCATG 768
XX      218 CysLysGlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThr 237
XX      769 TGTAAAGGGCTCCCGCTCTGGGAGAGATTTCTAGAGATTCTCAGACCTGACCGCAGCT 828
XX      238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu 257
XX      829 GTCTGTCCCGGTGCTGTGCGCTGCGCAAGGGGCTGACCTGCTGACCTGTGCTGATGAG 888
XX      258 GlnCysAlaIaIaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
XX      889 CAGTGTCTCGCGGCTGACAGGGGCGCAAGACATGCTAGTGGCTGGCTGCTGCTGCTGCT 948
XX      278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
XX      949 AACACAGTGGCATCTGTGAGCTGCTGACCTGCGCCACCTGGTGCACCTCAACACAGACAG 1008
XX      298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
XX      1009 TTGTAGTCCATGCCCAATCCCGAGGGCGGTATATTCATTCGGCCGCTGTGTACTGCTCC 1068
XX      318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
XX      1069 TGTCCCTAACAACTACCTTCTACGAGCTGGAGATCTGCACCTCGTCTGCGCCCTGCAC 1128
XX      329 ----- 329
XX      1129 AACCAAGAGGTGACAGCAGAGATGAACACAGCGGTGTGAGAAATGTCACCAACCCCTGT 1188
XX      329 ----- 329
XX      1189 GCCCGAATGTGCTATGTGTGCGCATGTGACACTTGGAGAGTGAAGGCAAGTTACCAAGT 1248
XX      329 ----- 329
XX      1249 GCCAATATCCAGAGATTGTGCTGCTGCTGCAAGAAATCTTTGGAGCTTGTCTGCGG 1308
XX      329 ----- 329
XX      1309 GAGAGCTTTGATGGGAGCCAGCGCTCCAAACATGCCCCGCTCAGCAGAGACAGTCCAA 1368
XX      329 ----- 329
XX      1369 GTGTTGAGACTCTGGAAGAGATCACAGTTTACTATATCTACATGTCAGATGGCCGGACAG 1428
XX      329 ----- 329
XX      1429 CTGGCTGACCTCAGCGCTTCCAGAAACCTGCAAGTAATCCGGGAGCAATTTGCACAAT 1488
XX      329 ----- 329
XX      1489 GGGCGCTACTCGCTGACCTTGAAGGGCTGGGATGACGTGGCTGGGGCTGCGCTACTG 1548
XX      329 ----- 329
XX      1549 AGGGAAGTGGGAGTGGAGTGGCTGCTATCCACATTAACACCCACCTCTGCTGCTGAC 1608

```

QY	329	----	329
Db	1609	ACGGTGCCTGGGACCAAGCTCTTTCGGAAACCGGACCAAGCTCTGCTCCACACTGCCAAC	1668
QY	329	----	329
Db	1669	CGGCCAGAGACAGATGTGTGGGCGAGGGGCTGGCCTGCCACACAGCTGTGGGCGCCGAGG	1728
QY	329	----	329
Db	1729	CAGTGTGGGCTCCAGGGGCCCCACCAGTGTGTCACTGACAGCAGTCTTCGGGGCCAG	1788
QY	329	----	329
Db	1789	GAGTGCCTGGAGAAATGCCAGTACTGTGCAGAGGGCTCCCGAGGAGTATGTGAATGCCAG	1848
QY	329	----	329
Db	1849	CAGTGTGGCTGCCACCCCTGAGTGTCAAGCCCAAGATGGCTAGTACCTGTTTGA	1908
QY	329	----	329
Db	1909	CCGGAGCTGACCAAGTGTGTGGCTGTGCCACTAATAAGACCCTTCCTTGCGCTGCC	1968
QY	329	----	329
Db	1969	CGGTGCCCAAGCGGTGTGAACCTGACCTCTCTACATGCCCCATCGGAAGTTCCAGAT	2028
QY	329	----	329
Db	2029	GAGAGGGCGATGCCAGCCTTCCCCCACTGACACCACCTCCTGTGTGAGACTGGAT	2088
QY	329	----	329
Db	2089	GACAAGGCTGCCCCGCCGAGAGAGAGCGACGCTCTGAGTCCATCATCTCTCGGCTG	2148
QY	329	----	329
Db	2149	GTTGGCATTTCTGTGTGTGTCTTGGGGGTGTCCTTGGGATCTCATCAAGCGACGG	2208
QY	329	----	329
Db	2209	CAGCAGAAGATCCGGAAAGTACAGATGCGGAGACTCTGCAGAAAGAGAGCTGGTGAG	2268
QY	329	----	329
Db	2269	CCGCTGACACTACGGAGCGATGCCCAACAGGCGCAAGTCCGATCCTGAAGAAGACG	2328
QY	329	----	329
Db	2329	GAGCTGAGGAAGGTGAAGTGTTCGATCTGGCGCTTTTGGCACAGTCTACAGGGGCATC	2388
QY	329	----	329
Db	2389	TGATATCCCTGATGGGAGAAATGTGAATAATTCAGTGGCCATCAAAAGTGTGAGGGAAC	2448
QY	329	----	329
Db	2449	ACATCCCCCAAGCAACAAGAAATCTTAGAGGAAGCATAGTGTGTGTGTGGC	2508
QY	329	----	329
Db	2509	TCCCATATATGTCCCGCCTTTTGGGCAATGCTGCTGACATCCACGGTGCAGCTGCTGACA	2568
QY	329	----	329
Db	2569	CAGTATATGCCATATGTGCTGCTCTTAGACATGTCGGGAAACCGGGAGCGCTGGGC	2628
QY	329	----	329
Db	2629	TCCAGAGCCTGCTGAATCTGTATATGACAGATTGCCAAGGGATAGCTACTGGAGGAT	2688

OY	329	-----		329
Dd	2689	GTCGGGCTGTACACAGGACTTGGCCGCTCGGAACGTCTGTGTCAAGATGCCAACCAT	2748	
OY	329	-----		329
Dd	2749	GTCAAATTACAGACTTCGGGGCTGGGCTGGCTGTGCACATTGAGAGACAGATACCAT	2808	
OY	329	-----		329
Dd	2809	GCAGATGGGGCAGAAGTGCCCATCAAGTGAGTGGCGGTGAGTCCATTCTCCGCGCGG	2868	
OY	329	-----		329
Dd	2859	TTTACCACACAGACTGATGTGTGGAGTTATGTTGCTGTGACTGTGTGGAGCTGATGACTTTT	2928	
OY	329	-----		329
Dd	2929	GGGGCCAACCTTACGATGGGATCCAGCCCCGGGAGATCCGTGACTGCTGAAAAAGGG	2988	
OY	329	-----		329
Dd	2989	GAGGGGCTGCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAAATGT	3048	
OY	329	-----		329
Dd	3049	TTCGATGATTCCTGTAATGTCCGCCAAGATTCCGGAGTTGGTGTCTGAATTTCTCCGC	3108	
OY	329	-----		329
Dd	3109	ATGGCAGAGGACCCCAGCCGCTTTGTGTGTCATCCAAATGAGAGACTTGGGCCACCACT	3168	
OY	329	-----		329
Dd	3169	CCCTTGACACACACTTCTACCGCTCACTGTGAGAGACGATGAGGGGAGCTGGTG	3228	
OY	330	-----	11	330
Dd	3229	CATGCTAGAGATTCCTGGTACCCACCAAGGCTTTCTTGTGCCAGACCTCCCGGGC	3288	
OY	331	AAGlglYmEtValInHshIArgHshIArgSerSerThrArgSeriYglYlAsp	350	
Dd	3289	GCTGGGGCAGTGTCCACACAGGACCGGAGCTCATTCACAGAGTGGCGTGGGAC	3348	
OY	331	LeuthrleuglYleugluProserGlugluAlaProargSerProeuAlaProser	370	
Dd	3349	CTGACACTAGGGCTGGAGCCCTCTGAAGAGAGGCCCCAGGCTCTCCACTGGCACCTCC	3408	
OY	371	GIUGlYlaaglYserAspValPheaspGlyAspLugLYmeGLyAlaIalYslYleu	390	
Dd	3409	GAAGGGCTGGCTCCGATGTAATTTGATGGTAGCTGGAAATGGGGCACCAAGGGGCTG	3468	
OY	391	GlnserLeuproThrHshAspProserProleuGlnArgTySerGluasproThrVal	410	
Dd	3469	CAAAAGCTCCCCACACATGACCCAGGCGCTTACACGGGTACAGTGAAGACCCACAATA	3528	
OY	411	ProleuproSerGluThrAspGlyTYrValAlaProLeuThrCyseSerProginProgu	430	
Dd	3529	CCCCGCCCTGTGAGTGTAGTGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGGCTGAA	3588	
OY	431	TYrValasnGlnProaspValAlatrgProGlnProProserProargGluYlProleuPro	450	
Dd	3589	TATGTGAACCGACCGATGTTCCGGCCCCCAGCCCCCTTCCCGAGAGGGCCCTCTGGCT	3648	
OY	451	AlaAlaIatrgProAlaGlyAlatrhrlleuGluAtrgAlaIalysTrhleuSerProglYlYasn	470	
Dd	3649	GCTGCCGACCTGCTGCTGCTGCACTGTGGAAAGGCCCAADACTCTCTCCCCAGGAAGAT	3708	
OY	471	GlyAlaValYlYasAspValPheAlaIarnegLgylAlaValaGlnAsnProglYlYleuthr	490	
Dd	3709	GGGGTCGCAAAAGACGTTTTTGGCTTTGGGGGTGGCGGTGAGAACCCGAGACTTGAACA	3768	
OY	491	ProGlnclYglYalaIalaproGlnProHshAspProIoalapheserProAlaIarheasp	510	

```

|||||
Db 3769 CCCGAGGAGAGCTGCCCTCCACAGCCCTCTCTGCTTCAGCCAGCCTTGAC 3828
Oy 511 AsnLeuTyTrpPaspGlnAspProGlnuArgGlyAlaProPserThPhyls 530
Db 3829 AACCTCTATTACTGAGACAGACCCACAGCGGGGCTCCACCCAGCCTTCAAA 3888
Oy 531 GlyThrProThralaGlnAspProGluTyrLeuGlyLeuAspValProAlaAlaPro 550
Db 3889 GGGACACTAGCGGAGAGAACCCAGACTACTGGCTGTGACGTGCCA-----GTGTGA 3942
Oy 551 AlaArgSerProSerPro 556
Db 3943 ACCGAGAGGCCAAGCTCG 3960

RESULT 15
AAF24297
ID AAF24297 standard; DNA: 9274 BP.
XX
AC AAF24297;
XX
DT 03-APR-2001 (first entry)
XX
DE HER2 transgene plasmid construct.
XX
KW Human; HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;
XX antibody; ds.
XX
OS Homo sapiens.
XX
PN WO200100244-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-US17229.
XX
PR 25-JUN-1999; 9905-0141316.
XX
PR 16-MAR-2000; 2000US-0189844.
XX
PA (GETH ) GENENTECH INC.
XX
PI Erickson S, Schwall R;
XX
DR WPI; 2001-061962/07.
XX
DR P-PSDB; AAB60167.
XX
PT Treating tumors, particularly breast cancers, which overexpress an ErbB
PT receptor and does not respond to an anti-ErbB antibody, comprises
PT conjugating the antibody to a maytansinoid -
XX
XX
PS Example 3; Fig 4; 92pp; English.
XX
CC The present invention provides a method of treating cancer by
CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In
CC particular, the antibody is directed against ErbB2 (also known as HER2
CC and p185neu). The method is particularly useful in the treatment of
CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,
CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.
XX
XX
SQ Sequence 9274 BP; 2152 A; 2539 C; 2488 G; 2095 T; 0 other;

Alignment Scores:
Pred. NO.: 3.58e-97 Length: 9274
Score: 2415.00 Matches: 527
Percent Similarity: 41.04% Conservative: 9
Best Local Similarity: 40.35% Mismatches: 32
Query Match: 64.59% Indels: 738
DB: 22 Gaps: 6

US-09-821-883-2 (1-690) x AAF24297 (1-9274)
Oy 2 ArgAlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 21

```

```

|||||
Db 1695 CGCCCCCGCCCTCCAGCCGGGTCCAGCCGAGACCATGAGAGTGGCG-----GCC 1745
Oy 22 LeuPhePheTrpLeuAspArgSerValLeuAlaLysLeuLeuAlaArgGlyAlaAlaSer 41
Db 1746 TTGGCGCGTGG-----GGGCTCTCTCTGCGCTCTTGGCCCCCGGAGCGGAGC 1796
Oy 42 ThrGlnAlaGlyThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThnIs 61
Db 1797 ACCCAATGTGACACCGGACAGACATGACATGCGGGTCTCCCTGCAATCCCGAGACCCAC 1896
Oy 62 LeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnLysLeuLeu 81
Db 1857 CTGGACATGCTCCCGCCACCTCTACACAGGCTGCGACGTGGTGCAGGGAAACCTGGAACCT 1916
Oy 82 ThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGlnGly 101
Db 1917 ACCTACTGCGCCCAATGACAGCTGTCTCTCTGAGGATTCAGAGGTGCAGGGC 1976
Oy 102 TyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuAlaVal 121
Db 1977 TACGTGCTCATCGCTCACACCAAGTGAGGAGTCCCACTGCAGAGGCTGGCATTTGTG 2036
Oy 122 ArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspPro 141
Db 2037 CGAGGACCCAGCTCTTGGAGACACTATGCCCTGGCCGTGTAGCAATGGAGACCCG 2096
Oy 142 LeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeuGlnLeu 161
Db 2097 CTGACAAATACACACCCCTGTACAGGGGCTCCCAAGAGAGCCCTGGGAGCTGCAGCTT 2156
Oy 162 ArgSerLeuThrGluLeuLeuLysGlyValLeuIleGlnArgAspProGlnLeuCys 181
Db 2157 CGAAGCTCTACAGAGATCTTGAAGAGGGGTCTTGTACAGCGGAACCCCAAGCTGTGC 2216
Oy 182 TyrGlnAspThrIleLeuTyrPylAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201
Db 2217 TACCAAGACACAGATTGTGGAAGACATCTTCCACAAGAACCAACGCTGGCTCTACA 2276
Oy 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysLysGlySer 221
Db 2277 CTGATACACACCAACCGCTCTCGGGCTGCACACCCCTGTCTCGATGTATAGGCTCC 2336
Oy 222 ArgCysTrpGlyLysSerSerGlnAspCysGlnSerLeuThrArgThrValCysAlaGly 241
Db 2337 CGCTGCTGGGGAGAGAGTCTGAGGATTTGCAGAGCTTCAGCGCCTGCTGTGGCGGT 2396
Oy 242 GlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAlaAla 261
Db 2397 GGGTGTCCCGCTGCAGAGGGGCCACTGCCACATGACTGTGCTGCCATGAGAGTGTGCTGC 2456
Oy 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281
Db 2457 GGGCGACAGGGGCCCAAGCACTGACTGCTGCTGCCCTGCCTCAACCAACAGAGGCG 2516
Oy 282 IleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMet 301
Db 2517 ATGTGTAGGTGACCTCCAGCCCTGTGACCAACACAGACAGCTTGTAGTGCATG 2576
Oy 302 ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn 321
Db 2577 CCNATCCGAGGGCGGTATACATTGGCGGCCAGCTGTGTGACTGCTGTCTCCCTACAC 2636
Oy 322 TyrLeuSerThrAspValGlySer----- 329
Db 2637 TACCTTTCTAGGAGACGTGGATCTGTGACACCTGTGTGCCCCCTGCACAAACAGAGGTG 2696
Oy 329 ----- 329
Db 2697 ACAGACAGAGATGAGAACACACAGCGGTGTGAGAAGTGACAGAACCCCTGTGCCGAGTGTGC 2756
Oy 329 ----- 329

```

Dh 2157 TATGCTGGGCAATGAGCACTTGGAGAGGTGAGGGAGTTACAGTGCATATATCCAG 2816  
Oy 329 ----- 329  
Db 2817 GAGTTTGTGGCTGCGAAGAAATCTTTGGAGCCTGGCATTTTCTGCCGAGAGCTTTGAT 2876  
Oy 329 ----- 329  
Db 2877 GGGAGCCAGCCTCCACACTGCCCCGCTCCAGCCAGAGCAGTCCAAGTGTGAGACT 2936  
Oy 329 ----- 329  
Db 2937 CTGGAAGATCACAGGTACTATACATCTCAGCATGGCCGAGACCTGCTGACCTC 2996  
Oy 329 ----- 329  
Db 2997 AGGCTCTTCCAGAACTGCAAGTATCCGGGAGCAATTTCTGCACAAATGGCGCTACTCG 3056  
Oy 329 ----- 329  
Db 3057 CTGACCCCTGCAAGGGCTGGGCATCAGCTGGGCTGGGCTCAGTGGGAACCTGGC 3116  
Oy 329 ----- 329  
Db 3117 AGTGGACTGGCCCTCATCCACATMACACCCACTCTGCTTCTGTCACAGGTGCTCCCTGG 3176  
Oy 329 ----- 329  
Db 3177 GACCAGCTCTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCCGCCAGAGGAC 3236  
Oy 329 ----- 329  
Db 3237 GAGTGTGGGCGAGGGCTGGGCTGCCACAGCTGTGCCGCCAGGGCACTGCTGGGGT 3296  
Oy 329 ----- 329  
Db 3297 CCAAGGCCCAACCACTGTCTCACTGACGCAAGTTCCTTGGGGCCAGAGTGGTGAG 3356  
Oy 329 ----- 329  
Db 3357 GAATGCCAGTACTGACAGGGGCTCCCAAGGAGTATGTGAATGCCAGGCACTGTTTGGCG 3416  
Oy 329 ----- 329  
Db 3417 TGCCACCCTGAGTGTACGCCCCAGAAATGGCTCAGTGAACGTGTTTGACCGGAGGCTGAC 3476  
Oy 329 ----- 329  
Db 3477 CAGTGTGGCTGTGCCCACTATAGAGACCCTTCTTGCCTGGCCGCTGCCAGC 3536  
Oy 329 ----- 329  
Db 3537 GGTGTAAACCTGACCTCTCTCATACATGCCATCTGGAAGTTTCCAGATGAGAGGCGCA 3596  
Oy 329 ----- 329  
Db 3597 TGCCAGCTTGCCCCATCACTGCAACCCACTCTGTGTGACCTGATGACAAAGGCTGC 3656  
Oy 329 ----- 329  
Db 3657 CCCGCGAGAGAGAGCCAGCCCTGTGACGTCCATGCTCTCTGCGGTGGTGGCATTTCTG 3716  
Oy 329 ----- 329  
Db 3717 CTGTGTGTGCTTTGGGGGTGTCTTTGGATCTCTCATCAAGCGACGCGACAGACATC 3776  
Oy 329 ----- 329  
Db 3777 CGAAGTACACGATCGGAGACTCTGTGAGGAACGAGAGTGTGGAGCCGCTGACACT 3836  
Oy 329 ----- 329  
Db 3837 AGCGAGCGATGCCCAACAGGCGCAGATGCGATCTCTGAAAGAGAGGAGCTGAGAG 3896

Oy 329 ----- 329  
Db 3897 GTGAAGTGTGTGATCTGGCGCTTTTGGCACACTTACAAGGCACTGTGATCCCTGAT 3956  
Oy 329 ----- 329  
Db 3957 GGGGAAATGTGAATAATTCAGTGGCCATCAAGTGTGAGGAAACACATCCCCAAA 4016  
Oy 329 ----- 329  
Db 4017 GCCAACAAGAAATCTTAGACGAAGCATACGTATGCTGTGTGGCTCCCATATGTC 4076  
Oy 329 ----- 329  
Db 4077 TCCGCTTCTGGCATCTGCTGACATCCAGGTGCAGCTGTGTGACACACACTTATGCC 4136  
Oy 329 ----- 329  
Db 4137 TATGCTGCTCTTAGACATGTCCGGGAAACCGGAGGCTGGGCTCCAGACACTG 4196  
Oy 329 ----- 329  
Db 4197 CTGAAGTGTGTATGACAGATTGCCAAGGGATGAGCTACCTGAGATGTGCGCTGTA 4256  
Oy 329 ----- 329  
Db 4257 CACAGGACTTGGCGCTCGGAACGTCTGTCAAGAGTCCCAACCATGTCAAAATTACA 4316  
Oy 329 ----- 329  
Db 4317 GACTTGGGCTGGCTCGGCTGCTGACATTGACGAGACAGATACATGAGATGGGGC 4376  
Oy 329 ----- 329  
Db 4377 AAGTGCCATCAAGTGAATGCGCTGAGTCCATCTCCGCGGCTTACACCACAG 4436  
Oy 329 ----- 329  
Db 4437 AGTATGTGTGAATTTATGTGTACTGTGTGAGAGCTGATGACTTTTGGGGCCAAACT 4496  
Oy 329 ----- 329  
Db 4497 TACATGGGATCCAGCCCGGAGATCCCTGACCTGTGMAAAGGGAGACGCTGCC 4556  
Oy 329 ----- 329  
Db 4557 CAGCCCCCATCTGCACCATTTGATCTACATGATCAATGTGATGATTTGAC 4616  
Oy 329 ----- 329  
Db 4617 TCTGAATGTGGCCCAAGATTCCGGAGTTGTGTCTGAATTTCTCCCATGCGAGGAC 4676  
Oy 329 ----- 329  
Db 4677 CCCAGGCTTTGTGTGATCCAGATGAGACTTGGGCCAGCCAGTCCCTTGACAGC 4736  
Oy 329 ----- 329  
Db 4737 ACCTTACCGCTCAGTCTGAGAGAGATGACATGGGGACCTGTGATGCTGAGAG 4796  
Oy 330 ----- 334  
Db 4797 TATCTGTACCCAGCAGGCTTTCTGTGTCACACCTGCCCGGGCGCTGGGGCATG 4856  
Oy 335 ValHisIshArgHisArgSerSerThrArgSerGlyGlyIyaPleuThrLeuGly 354  
Db 4857 GTCCACCAAGGACCCGACCTCATCTACCGAGGTGGCGGTGGGACTACACTGAGG 4916  
Oy 355 LeuGluProSerGluGluIyaIaProArgSerProLeuIaProSerGluGlyIyaGly 374  
Db 4917 CTGAGCCCTCTGAAGAGAGGCCCCAGGTCTCCTACCTGACACCTCCGAAGGGGCTGGC 4976





GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: May 8, 2003, 10:54:14 ; Search time 4337 Seconds

(without alignments)  
4630.144 Million cell updates/sec

Title: US-09-821-883-2  
Perfect score: 3739  
Sequence: 1 MRNAPLLLRASLSLGLF.....EPVQEGAPRRRAHHNNHH 690

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODEL=framer.p2n.model -DEV=alp  
-Q/cg2\_1/USFTO.spool/US09821883/runat\_28042003\_104749\_14566/app\_query.fasta\_1.839  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09821883 -GCN 1.1 3745 -runat\_28042003\_104749\_14566 -NCPU=6 -TCPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

## Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vl:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_pln:\*  
35: em.htg\_rtd:\*  
36: em.htg\_mam:\*  
37: em.htg\_vrt:\*  
38: em.sy:\*  
39: em.htgo\_hum:\*  
40: em.htgo\_mus:\*  
41: em.htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3739	100.0	2070	6	AX268286
2	3473.5	92.9	2091	6	AX268288
3	2974	79.5	1665	6	AX268285
4	2959.5	79.2	1692	6	AX268287
5	2423	64.8	4530	6	AR202597
6	2423	64.8	4530	6	AX282577
7	2423	64.8	4530	6	BD005474
8	2423	64.8	4530	6	121124
9	2423	64.8	4530	6	159745
10	2423	64.8	4530	6	HUMHR2A
11	2422	64.8	9274	6	AX060703
12	2418	64.7	4473	6	AR080259
13	2418	64.7	4473	6	AR167390
14	2418	64.7	4473	6	HSERB2R
15	2410.5	64.5	3768	6	AX060704
16	2410.5	64.5	3768	6	AX467229
17	2410.5	64.5	3768	6	AX481438
18	2405.5	64.3	3768	6	AR034479
19	2405.5	64.3	3768	6	AX201817
20	2405.5	64.3	3768	6	AX380923
21	2405.5	64.3	3768	6	AX384604
22	2405.5	64.3	3768	6	AX465456
23	2258.5	60.4	2385	6	AR082744
24	2258.5	60.4	2385	6	AR095963
25	2258.5	60.4	2385	6	AR143949
26	2188.5	58.5	1437	6	AX268289
27	2130	57.0	2781	6	AX380944
28	2129	56.9	2763	6	AX380942
29	2124.5	56.8	3780	4	AB008451
30	1968	52.6	3955	6	AX380924
31	1968	52.6	3955	6	121129
32	1968	52.6	3955	6	159750
33	1968	52.6	3955	10	RNNEUR
34	1961	52.4	3771	6	AX189662
35	1961	52.4	3771	6	AX380925
36	1955	52.3	4062	10	HANNEU
37	1680.5	44.9	2871	6	AX023363
38	1647	44.0	1872	6	AR181000
39	1613	43.1	1316	9	AF177761
40	1587	42.4	867	6	AX268303
41	1199.5	32.1	1191	6	AX268307
42	1188	31.8	1806	6	AX384608
43	1183	31.6	1755	6	AX384609
44	1183	31.6	1767	6	AX384607
45	1183	31.6	1773	6	AX384610

RESULT 1

## ALIGNMENTS

AX268286  
 LOCUS AX268286 2070 bp DNA linear PAT 29-OCT-2001  
 DEFINITION Sequence 7 from Patent WO0174855.  
 ACCESSION AX268286  
 VERSION AX268286.1 GI:16541538  
 KEYWORDS  
 SOURCE synthetic construct.  
 ORGANISM synthetic construct.  
 REFERENCE 1  
 AUTHORS Leus, R., Vidovic, D. and Graddis, T.  
 TITLE Compositions and methods for dendritic cell-based immunotherapy  
 JOURNAL Patent: WO 0174855-A 7 11-OCT-2001;  
 DENDREON CORPORATION (US)  
 FEATURES  
 source  
 1. 2070  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="HER500-HM-CSF construct"  
 BASE COUNT 425 a 707 c 568 g 370 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,18e-129 Length: 2070  
 Score: 3739.00 Matches: 690  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 US-09-821-883-2 (1-690) x AX268286 (1-2070)  
 QY 1 MetcrgAlaAlaProleuLeuLeuAlaAlaAlaSerleuSerleuGlyPheLeuPhe 20  
 Db 1 ATGAGACCTGACACCCCTCTCTGCGCAGGAGCAAGCCTTACCTTGCTTCTT 60  
 QY 21 LeuLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 40  
 Db 61 CTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 120  
 QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60  
 Db 121 TCGAGCCCAAGTGTGACAGGCGCAGACATGAGTGGCTCCTGCCAGTCCGGAACC 180  
 QY 61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnGlnLysnLeuGlu 80  
 Db 181 CACCTGGACATGCTCCGCCACCTCTACAGGGCTGCGAGGTGTCAGGAGAACTGGAA 240  
 QY 81 LeuThrTyrlleuProThrAsnAlaSerleuSerPheLeuGlnAspIleGlnIuValGln 100  
 Db 241 CTGACCTACCTGCGCCCAATGCGAGCTCTCTCTGAGGATATCCAGAGGTGTCAG 300  
 QY 101 G1yTyrlValLeuIleAlaHisAsnGlnValArgGlnValProleuGlnArgLeuArgIle 120  
 Db 301 GGGTACGTGCTCATCGCTCAACAACAGTAGGAGGTCCCACTGCGAGGCTGCGGATT 360  
 QY 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyrlAlaLeuAlaValLeuAspAsnGlyAsp 140  
 Db 361 GTGCGAGGACCCAGCTCTTGGAGACACTATGCTGCGCTGCTTGAACAATGAGAC 420  
 QY 141 ProleuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlyLeuGln 160  
 Db 421 CCGCTGAAACAATACACCCCTGTACAGGGGCTCCCAAGAGGCTGCGGAGACTGCAg 480  
 QY 161 LeuArgSerLeuThrGlnIleLeuGlyGlyValValLeuIleGlnArgAsnProGlnLeu 180  
 Db 481 CTTCGAAGCCTCAAGAGATCTTGAAGGAGGGGTCTTATCCAGCGGAACCCCACTC 540  
 QY 181 CysTyrlGlnAspThrIleLeuTyrlPylAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200  
 Db 541 TGTCTACAGACACAGATTTGTGGAAGGACATCTTCCACAAGAAACACAGCTGCTCTC 600  
 QY 201 ThrGlnIleAspThrAsnArgSerArgAlaCysHisPheProCysSerProMetCysLysGly 220

Db 601 ACACGTGATGAGACACCAACCGCTCTGCGGCTGCGCACCCCTGTTCTCGAGTGTAAAGGC 660  
 QY 221 SerArgCysTyrlGlyIleSerSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240  
 Db 661 TCCCGCTGCTGGGAGAGAGTGTGAGAGATTGTACAGACCTGACGCCACTGTCTGTGCC 720  
 QY 241 GlyIleCysAlaArgCysLysGlyProleuProThrAspCysHisGlnCysAla 260  
 Db 721 GGTGCTGTGCGCCCTGCAAGGGGCGCACCTGCGCTGCTGCTGCTGCTGCTGCTGCT 780  
 QY 261 AlaIleCysThrGlyProIleHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280  
 Db 781 GCCGCTGACAGGGGCCCAAGCACTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 QY 281 GlyIleCysGlnLeuHisCysProAlaLeuValThrTyrlAsnThrAspThrPheGlySer 300  
 Db 841 GGCATCTGTGAGCTGACCTGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 301 MetProAsnProGluIleArgTyrlThrPheGlyAlaSerCysValThrAlaCysProTyrl 320  
 Db 901 ATGCCCAATCCCGAGGCGCGGTATACATTCGGCGCCAGCTGTGACTGCTGCTGCTGCT 960  
 QY 321 AsnTyrlleuSerThrAspValGlySerGlyValAlaGlyIleMetValHisArgHisArg 340  
 Db 961 AACTACCTTTCTACGAGCTGAGTGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1020  
 QY 341 SerSerSerThrArgSerSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGlyGlu 360  
 Db 1021 AGCTCATCTACAGAGATGGGGGTGGGAGCTGACACTAGGGCTGAGAGCCCTCGAAGAG 1080  
 QY 361 GlnAlaProArgSerProleuAlaProSerGlyValAlaGlySerAspValPheAspGly 380  
 Db 1081 GAGCGCCCAAGTGTCTCCACTGCGCACCTCGGAAGGGGCTGCGCTCGATTTGATGAGT 1140  
 QY 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerleuProThrHisAspProSerPro 400  
 Db 1141 GACCTGGGAATGGGGGAGCGCAAGGGGCTGCAAGGCTCCCAACACATGACCCAGGCT 1200  
 QY 401 LeuGlnArgTyrlSerGlnAspProThrValProleuProSerGlyThrAspGlyTyrlVal 420  
 Db 1201 CTACAGGGGTACAGTAGAGACCCCAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
 QY 421 AlaProleuThrCysSerProGlnProGluTyrlValAsnGlnProAspValArgProGln 440  
 Db 1261 GCCCCCTGACCTGACCCCGCCCAAGTATGTGAACCAAGCAGATGTTGCGGCCAG 1320  
 QY 441 ProProSerProArgGlyGlyProleuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460  
 Db 1321 CCCCCTTCCGCCGAGAGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
 QY 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480  
 Db 1381 AGGCGCAAGACTCTCTCCAGGGAAGATGGGGCTCTCAAAAGCGTTTGTGCTTGGG 1440  
 QY 481 GlyAlaValGlnAsnProGluTyrlLeuThrProGlnGlyAlaAlaProGlnProHis 500  
 Db 1441 GGTGCTGTGAAGAACCCCGAGTACTTGACACCCAGAGAGAGGTGCGCCCTCAAGCC 1500  
 QY 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyrlTyrlPaspGlnAspProPro 520  
 Db 1501 CTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560  
 QY 521 GlnArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGlnAsnProGluTyrl 540  
 Db 1561 GAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
 QY 541 LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSerThrGlnPro 560  
 Db 1621 CTGGGTCTGAGAGCTGCGAGGGCGCCAGCCCGCTGCGCCACCCACACACACCC 1680  
 QY 561 TrpGlnHisValAsnAlaIleGlnGlnAlaArgArgLeuAsnLeuSerArgAspThr 580



QY 432 ValaIngInProAspValaArgProGInProProSerProArgGluGlyProLeuProAla 451  
 DB 1321 GTGAACAGCCAGATGTTGGGGCCAGCCCTTGCCCGAGAGGGGCTCTGCTCT 1380  
 QY 452 AlaArgProAlaGlyAlaThrLeuGlnArgAlaLysThrLeuSerPheGlyLysAsnGly 471  
 DB 1391 GCGCCACCTGGTGGTCCACCTCTGMAAGGGCCAAAGACTCTCTCCCGAGGAAAGATGG 1440  
 QY 472 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluThrLeuPro 491  
 DB 1441 GTGCGCAAGACGTTTGGCTTGGGGGTGGCGTGGAGAACCCGAGTACTTGACACC 1500  
 QY 492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511  
 DB 1501 CAGGAGAGAGCTGCCCTCAGCCGCCAGCCCTCTCTGCTTCAAGCCAGCTTGACAAAC 1560  
 QY 512 LeuThrTrpThrPheArgLysAspProGluArgGlyAlaProProSerThrPheLysGly 531  
 DB 1561 CTCTATTACTGGAGACAGGACCCACAGAGGGGGCTCCACCCAGACCTTCAAAAGG 1620  
 QY 532 ThrProThrAlaGluAsnProGluThrLeuGlyLeuAspValProAlaAlaAlaProAla 551  
 DB 1621 ACACCTAGGCGAGAAACCCAGAGACTGGGTCTGGACGTGCCAGCGGCCGCCACC 1680  
 QY 552 ArgSerProSerProSerThrGlnProTrpGlnHisValAsnAlaAlaGlnGluAlaArg 571  
 DB 1681 CGCTACCCCAACCTGCTACCCCGCCCTGAGAACATGTAGATGCCATCAAGAACCTCTG 1740  
 QY 572 ArgLeuAsnLeuSerPheArgAspThrAlaAlaGluMetAsnGluThrValGluValLe 591  
 DB 1741 AGCCCTCAATATGACATCGCTCTGTGAGAACCAAAAGAGAGAGATGACATCATCTC 1800  
 QY 592 SerGluMetPheAspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuThrLys 611  
 DB 1801 TCTATATAGTTCTTCATCATCAGAGCGGCAGCATGTGTGACAGCCCGCTTAACCTATAC 1860  
 QY 612 GlnGlyLeuArgGlySerLeuThrLysLeuLysGlyProLeuThrMetMetAlaSerHis 631  
 DB 1861 CAGGCTTACAGGGGCAACCTCAACCAACTCAATGAGCGCTTGACCATGATGACGAC 1920  
 QY 632 TyrLysGlnHisCysProProThrProGluThrSerCysAlaThrGlnIleLeuThrPhe 651  
 DB 1921 TACCAAGCAACTGCTCCCAACCCCGGAACTGACTGAATGAAAGTACACCACTTT 1980  
 QY 652 GluSerPheLysGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpGlu 671  
 DB 1981 GAGGATTCATAAAGAACCTTAAAGGCTTCTGTTGATATCCCTTTGACATGCTGGAG 2040  
 QY 672 ProValGlnGluGlyAlaProProProProAlaAlaAlaHisHisHisHisHis 690  
 DB 2041 CCGGTCCAGAAAGGGCGCCACCCCGCCG-----GGCATTCACATCACCATCAC 2091  
 RESULT 3  
 AX268285 1665 bp DNA linear PAT 29-OCT-2001  
 LOCUS AX268285  
 DEFINITION Sequence 6 from Patent WO0174855.  
 ACCESSION AX268285  
 VERSION AX268285.1 GI:16541537  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 synthetic construct.  
 artificial sequences.  
 REFERENCE  
 1 Laus R., Vidovic D. and Graddis T.  
 AUTHORS Compositions and methods for dendritic cell-based immunotherapy  
 TITLE Patent: WO 0174855-A 6 11-OCT-2001;  
 JOURNAL DENDREON CORPORATION (US)  
 FEATURES  
 location/Qualifiers  
 1..1665  
 source  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="HER500 construct"

BASE COUNT 330 a 568 c 465 g 302 t  
 ORIGIN  
 Alignment Scores:  
 Score: 4.05e-101 Length: 1665  
 Pred. No.: 2974.00 Matches: 534  
 Percent Similarity: 99.46% Conservative: 0  
 Best Local Similarity: 99.46% Mismatches: 1  
 Query Match: 79.54% Indels: 3  
 DB: 6 Gaps: 1  
 US-09-821-883-2 (1-690) x AX268285 (1-1665)  
 QY 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20  
 DB 1 ATGAGAGCTGACCCCTCTCTCTGCGCCAGGCGACAGAACCTTACGCTTGGCTTGT 60  
 QY 21 LeuLeuPhePheThrPheAspArgSerValLeuAlaLysGlyLeuAlaArgGlyAlaAla 40  
 DB 61 CTGCTTTTCTGTGGCTAGCCAGAGTGTACTAGCCAAAGAGTTGGCGGGGGCCGCG 120  
 QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60  
 DB 121 TCGACCCAGTGTGACCCGCGCAGACATGAAGCTGGCGCTCCCTGCCAGTCCGAGACC 180  
 QY 61 HisLeuAspMetLeuArgHisLeuThrGlnLysCysGlnValValGlnGlyAsnLeuGlu 80  
 DB 181 CACCTGAGCATGTCTCCGCGCACCTCTACACAGGGGTGCCAGGTGGTGCAGAAACCTGGA 240  
 QY 81 LeuThrTrpLeuProThrAspAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100  
 DB 241 CTCACCTACCTGCGCCACCAATGCCAGCTGTCTTCCGAGATATCCAGAGAGTGCAG 300  
 QY 101 GlyThrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgLle 120  
 DB 301 GGCTACGTCGTCTCATCTGCTCACAACCAATGAGCAGGTCCCACTGCCAGAGCTCGGATT 360  
 QY 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140  
 DB 361 GTGGAGAGCACCAGCTCTTTGAGAGACAACTATGCCCTGGCGGCTGTGACAAATGGAGAC 420  
 QY 141 ProLeuAsnAsnThrThrProValThrGlyLysAspProGluGlyLeuArgGlnLeuGln 160  
 DB 421 CCGGTGAACAATACCACCCCTGTCAACAGGGGCTCCCGAGAGGCTGGGGACTGTGAG 480  
 QY 161 LeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAspProGlnLeu 180  
 DB 481 CTTCGAACCTCACAGAGATCTTGAAGAGAGGGCTTGATGCCAGCGGAACCCAGCTC 540  
 QY 181 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200  
 DB 541 TGTACACAGGACAGATTTGTGGAGAGACATCTTCCACAAGAACCAACACCTGGCTCTC 600  
 QY 201 ThrLeuIleAspThrAspArgSerArgAlaCysHisProCysSerProMetCysLysGly 220  
 DB 601 ACCTGATAGACACCAACCGCTCTGGGCTGCCACCCCTCTCTCCGATGTGAAGGCG 660  
 QY 221 SerArgCysTrpGlyGluSerSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240  
 DB 661 TCCCGCTCTGGGAGAGAGTTGTGAGGATTTGTACAGACCTGAGCGGACAGCTGTGCC 720  
 QY 241 GlyLysCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 260  
 DB 721 GGTGGCTGTGCCCTGCAAGGGGCACTGCCACTGACTCTCCATGACACAGTGTGCT 780  
 QY 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280  
 DB 781 GCGGAGCTGACAGGCGCCCAACACTGTGACCTGTGGCTGTCCCTTCAACACACAGT 840  
 QY 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGluSer 300  
 DB 841 GGCACTGTGTAGCTGCACTGCCAGCCCTGTGACCTACATCAACACAGACACGTTTGAGATCC 900

```

QY 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
DB 901 ATGCCCAATCCGAGGCGGATATGATTCGGCGCCACCTGTGTACAGCCGTCCCTAC 960
QY 321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHisArgHisArg 340
DB 961 AACCTACCTTTTACCGAGCTGGAGATCGGGCGCTGGGGCATGTGTCCACACAGCACCGC 1020
QY 341 SerSerSerThrArgSerGlyGlyValAspLeuThrLeuGlyLeuGluProSerGluGlu 360
DB 1021 AGCTATCTACAGAGATGGCGGTGGAGCTGACACTAGGGCTGGAGCCCTTAAAGG 1080
QY 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
DB 1081 GAGGCCCCAGGTCTCCACTGGCACCTCCGAAAGGGCGGTGGCTCGAGATATTATGAT 1140
QY 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
DB 1141 GACCTGGGAATGGGGGCGAGCCAAAGGGCTGCMAAGCCCTCCACACATGACCCAGCCCT 1200
QY 401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420
DB 1201 CTACAGCGGTACAGTGAAGACCCACAGTACCCCTGCTGTGAGACTGATGGTACGTT 1260
QY 421 AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440
DB 1261 GCGCCCTGACCTGGACGCCCGCCTGATATGTGAACACCCAGATGTTGGGCCAG 1320
QY 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGln 460
DB 1321 CCCCCTTGCCCGGAGAGGGCCCTGCTGCTGCCGACCCGCTGGGCGCCACTGTGAA 1380
QY 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
DB 1381 AAGGCAAGACTCTCTCCCAAGAGATGGGTGCTCAAGAGCTTTTTCCTTTGGG 1440
QY 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHis 500
DB 1441 GGTGCGGTGAGAAACCCGAGTACTGACACCCGAGGAGGAGCTGCCCTGAGCCCA 1500
QY 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro 520
DB 1501 CCTCCTCCTGCTTACGCCACCTTCGACAACTCTATTACTGGAGCACAGACCCACCA 1560
QY 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 540
DB 1561 GAGCGGGGGGCTCCACCCAGCACTTCAAGGACACCTACGGCAGAGAACCCAGATAC 1620
QY 541 LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSer 557
DB 1621 CTGGGCTTGAGACGTGCGAGCGGCC-----GCACA-TCACCATCAACATCA 1664

```

RESULT 4  
 AX268287 1692 bp DNA linear PAT 29-OCT-2001  
 LOCUS Sequence 8 from Patent W00174855.  
 DEFINITION AX268287  
 ACCESSION AX268287  
 VERSION AX268287.1 GI:16541539  
 KEYWORDS  
 ORGANISM synthetic construct.  
 SOURCE synthetic construct.  
 REFERENCE 1  
 AUTHORS Laus, R., Vidovic, D. and Graddis, T.  
 TITLE Compositions and methods for dendritic cell-based immunotherapy  
 JOURNAL Patent: WO 0174855-A 8 11-OCT-2001;  
 DENDREON CORPORATION (US)  
 FEATURES  
 location/Qualifiers  
 1..1692  
 /organism="synthetic construct"  
 /db\_xref="taxon:32630"  
 /note="HER500\* construct"  
 BASE COUNT 338 a 573 c 470 g 311 t

```

ORIGIN
Alignment Scores:
Pred. No.: 1,4e-100
Score: 2959.50
Percent Similarity: 97.88%
Best Local Similarity: 97.88%
Query Match: 79.15%
DB: 6 Gaps: 2
US-09-821-883-2 (1-690) x AX268287 (1-1692)
QY 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
DB 1 ATGACAGCTGACACCCCTCTCTCTGCGCAGGCGAGCAACCTTAGCTTGCTTGT 60
QY 21 LeuLeuPhePheTyrPheAspArgSerValLeuAlaLysGlyLeuAlaArgGlyAlaAla 40
DB 61 CTGCTTTTCTGCTGCTGACCGAAGTGTACTAGCCAAAGAGTTGGCGGGCGCGC 120
QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
DB 121 TCGACCCAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCGACATCCGAGACC 180
QY 61 HisLeuAspMetLeuAlaArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
DB 181 CACCTGGACATGTCTCCGCCACTCTACACAGGGCTCCAGGTGTGTCCAGGAAACCTGGAA 240
QY 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIleValGln 100
DB 241 CTCACCTACCTGCCCCACCAATCCAGCTGCTTCCTGACAGATATCCAGAGAGTGCAG 300
QY 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
DB 301 GGTACGTGTCTATGTCTACAAACCAAGTGAAGAGGTCCCACTCCAGAGAGTGGGAT 360
QY 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
DB 361 GTGGAGGACCCACACTCTTTTAGAGACAACTATGCCCTGGCTGTACAACTGAGAGC 420
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
DB 421 CCGGTGAACAATACCAACCCCTGTACAGAGGGCTCCCAAGAGGCTCCGGAAGCTGCAG 480
QY 161 LeuArgSerLeuThrGlnIleLeuLysGlyValValLeuIleGlnArgAspProIleu 180
DB 481 CTTCGAGGCTTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCAAGCTC 540
QY 181 CysTyrGlnAspThrIleLeuThrPheLysAspIlePheHisLysAspAsnGlnLeuAlaLeu 200
DB 541 TGTACACAGACAGATTTTGTGGAAGACATCTTCCACAAACACAGCTGGCTCTC 600
QY 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
DB 601 ACATGATAGACACCAACCGCTCTGGGCTGCCACCCCTGTTCGATGTGAAGGC 660
QY 221 SerArgCysTyrPheGlyLysSerSerGlnAspCysGlnSerLeuThrArgThrValLysAla 240
DB 661 TCCGCTCTCTGGGAGAGATTTGAGGATTTCTGAGGCTGTACGCGCACTGTCTGTGCTCC 720
QY 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrArgPheCysCysHisGluGlnCysAla 260
DB 721 GGTGGCTGTGGCGCTGCAAGGGGCACTGCCCACTGACTGTGSCATGAGAGAGTGTCT 780
QY 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
DB 781 GCGGCTGACAGGGGCCCCAAGACACTGACTGCGCTGGCTGCTCCACTTCAACACAGT 840
QY 281 GlyIleCysGluLeuHisCysArgProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
DB 841 GGCATGTGTGACTGTACTGCGCCAGCCCTGGTGCACCTTCAACACAGACAGCTTTAAGTCC 900
QY 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysArgProTyr 320

```

```

DB 901 ATGCCCATCCGAGGCGCGTATTCATTCCGCCAGCTGTGTACCTGCTCCCTAC
OY 321 AsnTYrLeuSerThrAspValGlySer-----GlyAla 331
DB 961 AACCTACCTTTTACGAGAGTGAGATCCGCTACCATCATTAATTTCGAGAGAGTTGGCCCT
OY 332 GlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyGlyLeu 351
DB 1021 GGGGGCATGTGTCACACAGGACCGCAGCTCATCTACAGAGAGTGGGGTGGGACCTG
OY 352 ThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGlu 371
DB 1081 ACACTAGGGCTGAGACCTCTGTAAAGAGAGGCCCCAGAGTCTCCACTGGACCTCCGAA
OY 372 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyGlyLeu 391
DB 1141 GGGGCTGCTCCGATGATTGATGTGTCACCTGGGAATGGGGCAGCCAGGGCGCTGCA
OY 392 SerLeuProThrHisAspProSerProLeuGlnArgTYrSerGlyAspProThrValPro 411
DB 1201 AGCTTCCTCCACATGAGCCACGCTCTACAGCGGTACAGAGAGAGCCACACTACCC
OY 412 LeuProSerGluThrAspGlyTYrValAlaProLeuThrCysSerProGluProGluTYr 431
DB 1261 CTGCCCTCTGAGACTGATGGTACGTTGCCCTCCGACCTGACGCCCCACGCTCAATAT
OY 432 ValAsnGlnProAspValArgProGlnProProSerProArgGlyGlyProLeuProAla 451
DB 1321 GTGAACCGACGAGATGTGGCGCCCGCCCTTCGCCCCAGAGAGGGCCCTGCTGCTCT
OY 452 AlaArgProAlaGlyAlaThrLeuGluArgAlaGlyThrLeuSerProGlyGlyAsnGly 471
DB 1381 GCCGACCTGCTGGTGGTCCAGCTGGAAAGGCGCAAGACTCTCTCCCGGAGAAAGATGG
OY 472 ValValGlyAspValPheAlaPheGlyGlyAlaValGluAsnProGluTYrLeuThrPro 491
DB 1441 GTGCGAAGAGCTTTTGGCTTTGGGGGTGCCGTGGAGAACCCGCACTACTTGACACCC
OY 492 GlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsn 511
DB 1501 CAGGAGAGAGTGGCCCTCCACCCACCCCTCTCTGCTTCAGGCCGCTTCGACACAC
OY 512 LeuTYrTYrPASPGLAspProProGluArgGlyAlaProProSerThrPheGlyGly 531
DB 1561 CTCGATTACTGGACACGAGACCCACAGAGCGGGGCTCCACCCAGCACCTTCAAAGGG
OY 532 ThrProThrAlaGluAsnProGluTYrLeuGlyLeuAspValProAlaAlaAlaProAla 551
DB 1621 ACACCTACGCGAGAGAACCCAGAGTACCTGGTCTGGACGTGCCAGCGGGCC-----GCA 1674
OY 552 ArgSerProSerProSer 557
DB 1675 CA-TCACCATCACCATCA 1691

RESULT 5
AR202597 4530 bp DNA linear PAT 20-JUN-2002
LOCUS AR202597 Sequence 4 from patent US 6365151.
DEFINITION AR202597
ACCESSION AR202597 GI:21498768
VERSION AR202597.1
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 4530)
AUTHORS
Halpern, M.S. and England, J.M.
TITLE
Cellular immunogens comprising cognate proto-oxogenes
JOURNAL
Patent: US 6365151-A 4 02-APR-2002;
FEATURES
Location/Qualifiers
1..4530
/organism="unknown"
BASE COUNT 922 a 1382 c 1346 g 880 t

```

```

ORIGIN
Alignment Scores:
Pred. No.: 1,49e-80 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.148 Conservative: 5
Best Local Similarity: 40.758 Mismatches: 21
Query Match: 64.80% Indels: 736
Gaps: 4
DB: 6

US-09-821-883-2 (1-690) x AR202597 (1-4530)
OY 2 ArgAlaAlaProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
DB 94 CGCGCCCTCCACAGCGGGGTCCAGCGGATGGGCGCGGAGCGGAGTGGACCATG 153
OY 18 PheLeuPheLeuLeuPhePheThrLeuAspArgSerValLeuAlaGlyGluLeuAlaArg 37
DB 154 GAGCTGGCGGCTGTGTGCGGCTGG-----GGGCTCTCTGCGCCCTCTTCCGCCCC 204
OY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetGlyLeuArgLeuProAlaSer 57
DB 205 GGAGCGCGGAGCAGCCAGATGTGACCGCGCACAGACATGAAAGCTGGGCTCCCTGCCAGT 264
OY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTYrGlnGlyCysGlnValGlnGly 77
DB 265 CCGGAGACCCACCTGGACATGCTCCGCGCACCTTACACAGGGCTGCGACGTGTGTCAGGGA 324
OY 78 AsnLeuGluLeuThrTYrLeuProThrAsnAlaSerLeuSerPheLeuGluAlaPheGln 97
DB 325 AACCTGGAACTACCTTACCTCCACCAATGCCAGCTGTCTTCTGCAAGATATCCAG 384
OY 98 GluValGlnGlyTYrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
DB 385 GAGGTGACGAGGCTACGTCCTCATGCTCACAACCAAGAGGAGGTGCCACTGCAGAGG 444
OY 118 LeuArgGlyLeuValArgGlyThrGlnLeuPheGluAspAsnTYrAlaLeuAlaValLeuAsp 137
DB 445 CTGGGATGTGGCAGAGGACCCAGCTCTTTGAGGACCACTTGCGCTGCGCTCTACAC 509
OY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
DB 505 AATGGAGACCGCGTGAACAATACCAACCCCTGTACAGAGGGCTCCACAGAGGGCTGGG 564
OY 158 GluLeuGlnLeuArgSerLeuThrGluLeuLeuGlyGlyValLeuLeuGlnArgAsn 177
DB 565 GAGCTGACGCTTCGAAAGCTCTACAGAGATCTTGAAGAGGAGGGCTTGATGACACGGGAC 624
OY 178 ProGlnLeuCysTYrGlnAspThrThrLeuThrPheAspPheLeuHisLeuAsnGln 197
DB 625 CCCCAGCTCTCTCCAGAGACAGATTTGTGGAAAGGACATCTTCCACAGAGAACCCAG 684
OY 198 LeuAlaLeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
DB 685 CTGGCTGTACACATGATAGACACCAACCGGCTGCGGGCTCCACCCCTGTCTCCAGTG 744
OY 218 CysGlyGlySerArgCysTYrGlyGlyLeuSerSerGluAspCysGlnSerLeuThrArgTYr 237
DB 745 TGTAAAGGCTCTCCCTCTGCGGAGAGATTTCTAGAGATTGTCTAGACCTGACCCGAGCT 804
OY 238 ValCysAlaGlyGlyCysValArgCysValGlyProLeuProThrArgPheCysValGlu 257
DB 805 GTCTGTCCCGGTGGCTGTGCGGCTGCAAGGGGCGGACCTGCGGCTGCTGCTGCTGAG 864
OY 258 GlnCysAlaAlaGlyCysThrGlyProGluHisSerAspCysValAlaCysLeuHisPhe 277
DB 865 CAGGTGTCTGCGGCTGACAGGCGGCCCAAGACATCTGACCTGCGGCTGCTGCTGCTGCTG 924
OY 278 AsnHisSerGlyLeuCysGlyLeuHisCysProAlaLeuValThrTYrAsnThrAspThr 297
DB 925 AACCAAGTGGAGATCTGTAGCTGCTGACCTGCGGAGCTGCGGAGCTGACCAAGAGACAG 984
OY 298 PheGluSerMetProAsnProGluGlyArgTYrThrPheGlyAlaSerCysValThrAla 317

```

|||||  
Db 985 TTGAGTCATGCCCAATCCGAGGCCGCTATACATTGGCGCCAGCTGTGACTGCC 1044  
QY 318 CysProTyrAsnTyrIleuSerThrAspValIglySer----- 329  
Db 1045 TGTCCCTACACTACCTTCTACGGACGTGGATCTCTGCACCCTGCTGCCCCCTGCAC 1104  
QY 329 ----- 329  
Db 1105 AACCAAGAGTGACAGCAGAGATGSAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGT 1164  
QY 329 ----- 329  
Db 1165 GCCCAGTGTCTATGTCTGGGCATGAGACATTGCGAGAGGTGAGGGCAGTTACAGT 1224  
QY 329 ----- 329  
Db 1225 GCCAATATCCAGAGATTGTCTGGCTGCAGAAAGATCTTTGGAGCCTGSCATTTCTGCCG 1284  
QY 329 ----- 329  
Db 1285 GAGAGCTTTGATGGGAGCCAGCCCTCCAACTGCCCGCTCCAGCCAGACAGCTCCAA 1344  
QY 329 ----- 329  
Db 1345 GTGTTGAGACTCTGSAAGAGATCACAGGTACTATACATCTCAGCATGCGCGACAGC 1404  
QY 329 ----- 329  
Db 1405 GTGCTGACCTGAGCCTCTTCAGAACCTGCAGAAATATCCGGGGAGCAATTTCTGCACAT 1464  
QY 329 ----- 329  
Db 1465 GCGGCTACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTCGCTACTG 1524  
QY 329 ----- 329  
Db 1525 AGGGAAGTGGGAGTGAAGTGGCCCTCATCACATACACCACCTCTGCTTGGTGCAC 1584  
QY 329 ----- 329  
Db 1585 ACGGTGCCCTGGGACAGCTTTTGGGAACCGCACCAAGCTCTGCTCCACACTGCCAAC 1644  
QY 329 ----- 329  
Db 1645 CGGCCAGAGAGAGTGTGTGGGCGAGGGCTGGCCTGCACACAGCTGTGCCCGAGGG 1704  
QY 329 ----- 329  
Db 1705 CACTGCTGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTGGGGCCAG 1764  
QY 329 ----- 329  
Db 1765 GAGTGGCTGAGAGATGCCAGTACTGCAAGGGCTCCCGAGGGATATGTGAATGCCAGG 1824  
QY 329 ----- 329  
Db 1825 CACTGTTGCCGTGCCACCCTGAGTGTCAAGCCCAAGATGGCTCACTGACTGTTTGA 1884  
QY 329 ----- 329  
Db 1885 CCGAGAGCTGACCAAGTGTGTGGCTGTGCCACTATAGAGACCTCCCTTGTGCGTGGCC 1944  
QY 329 ----- 329  
Db 1945 CGCTGCCCAAGCGGTGTGAACCTGACTCTCTCTACATGCCCATTTGGAAATTTCCAGAT 2004  
QY 329 ----- 329  
Db 2005 GAGGAGGGGAGATGCCAGCCTTGCCCCATCACTGCACCCACTCCTGTGTGGACTGGAT 2064  
QY 329 ----- 329

Db 2065 GACAAAGGCTGCCCGCCGAGCAGAGACCAAGCCCTGTGACGTCCATCCTCTTCCGGTG 2124  
QY 329 ----- 329  
Db 2125 GTTGGCATTTCTCTGTGTGTGTCTTTGGGGGTGTCTTTGGGATCTCTATCAAGCAGG 2184  
QY 329 ----- 329  
Db 2185 CAGCAGAAGATCCGGAAATGACAGATGGCGAGACTGTCTCAGAGAAACGAGCTGGTGAG 2244  
QY 329 ----- 329  
Db 2245 CCGCTGACACTAGCGGAGCGATGCCAACAGCGCAGATCGGATCCTGAAAGAGAGC 2304  
QY 329 ----- 329  
Db 2305 GAGCTGAGAGAGTGAAGTGTCTTGGATCTGGCGCTTTTGGCACAGTCTCAAGGGCATC 2364  
QY 329 ----- 329  
Db 2365 TGGATCCTGATGGGAGAAATGTGAATAATTCAGTGGCCATCAAGTGTGAGGGAAAC 2424  
QY 329 ----- 329  
Db 2425 ACATCCCCCAAGCCAAACAAAGAAATCTTAGAGAAAGATAGTATGAGCTGTGTGGC 2484  
QY 329 ----- 329  
Db 2485 TCCCATATGTCTCCCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGTGACA 2544  
QY 329 ----- 329  
Db 2545 CAGCTTATGCCCTATGGCTGCCTTTAGACATGTGCCGGAAACCGGGAGCGCTGGGC 2604  
QY 329 ----- 329  
Db 2605 TCCAGAGACTGTCTGAAGTGTATGAGATTTGCCAAGGGATGAGTACTGGAGAT 2664  
QY 329 ----- 329  
Db 2665 GTGCGCTGTACACAGGACTTTGGCCGCTCGGAACGTCTGTCTCAAGATCCCAACAT 2724  
QY 329 ----- 329  
Db 2725 GTCAAAATTACAGACTTGGGCTGGCTGCTGTGACATTGACAGACAGATACAT 2784  
QY 329 ----- 329  
Db 2785 GCAGATGGGGGCAAGTGCATCAAGTGAATGGCGTGGATCCATTCTCCGCCGCGG 2844  
QY 329 ----- 329  
Db 2845 TTCACCCACACAGATGATGTGTGAGATTATGGTGTGACTGTGTGGAGCTGATGACTTT 2904  
QY 329 ----- 329  
Db 2905 GGGGCCAAACTTACAGTGGATCCAGCCCGGAGATCCTGACCTGCTGAAAAAGGG 2964  
QY 329 ----- 329  
Db 2965 GAGCGCTGCCAGCCCCCATCTGCACCATGTGATGTCTACATGATCATGTGTCAAATGT 3024  
QY 329 ----- 329  
Db 3025 TGGATGATTGACTGTGAATGTGTGGCCAAAGATTCCGGAGTGTGTGTGAATTTCTCCCG 3084  
QY 329 ----- 329  
Db 3085 ATGGCAGGAGCCCGCAGCGTTTGTGTATTCAGAAATGAGAGACTTGGGCCAGCAGT 3144  
QY 329 ----- 329  
Db 3145 CCCTTGGACAGACCTTCTACCGCTCACAGTGTGAGAGAGATGACATGGGGAGCCTGTG 3204





|||||  
Db 565 GAGGTGACAGCTTCGAAGGCTCAGAGATCTTGAAGAGGGGCTTGTATCCAGGGAAC 624  
Oy 178 ProGlnLeuCySTyTglnAspThrIleLeuTrpIysAspIlePheHisLysAsnAsnGln 197  
Db 625 CCCCAGCTCTGTCTACAGAGACAGATTTTGTGGAGGAGCATCTTCACAGAACCAACAG 684  
Oy 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPromet 217  
Db 685 CTGGCTCTACACTATATAGACACACACCGCTCTCGGGCTGCACCCCTGTTCTCCGATG 744  
Oy 218 CysLysGlySerArgCysTrpGlyIuSerSerGluAspCysGlnSerLeuThrArgThr 237  
Db 745 TGTAAAGGCTCCCGCTGCTGGGAGAGATTCTGAGGATTGACAGCTGACGCGGCACT 804  
Oy 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGln 257  
Db 805 GTCTGTGCGGTGGCTGTGGCCGCTGCAGAGGGCCACCTGCACCTGACTGCTCCATGAG 864  
Oy 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277  
Db 865 CAGTGTCTGCGCGGTGCAGAGGGCCCCCAAGCACCTGTGACTGCTGCTGCTCCACTTC 924  
Oy 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyraThrAspThr 297  
Db 925 AACACAGATGGCATCTGTAGAGCTGCACCTGCCAGCTGTGCTACACACAGACAG 984  
Oy 298 PheGluSerMetProAsnProGlnIuLysArgTyrrhrPheGlyAlaSerCysValThrAla 317  
Db 985 TTTGAGTCCATGCCCCCAATCCCGAGGGCGGTATACATTTCGGCGCAGCTGTGTGATGCC 1044  
Oy 318 CysProTyraSerTrpLeuSerThrAspValGlySer 329  
Db 1045 TGTCTCTACACTACTTTCTTACGAGACGTGGATCTCTGACCCCTGTCTGCCCTGCAC 1104  
Oy 329 329  
Db 1105 AACCAAGAGTGCACAGAGAGATGGAACACACGCGTGTGAGAAATGCAGAACCCCTGT 1164  
Oy 329 329  
Db 1165 GCCCAGTGTGCTATGCTGTGGGATGAGACCTTGCAGAGAGTGAGGGCAGTTACCAAT 1224  
Oy 329 329  
Db 1225 GCCAATATCCAGAGATTTGCTGCTGCAGAAAGATCTTTGGAGGCTGCAATTTCTGCCG 1284  
Oy 329 329  
Db 1285 GAGAGCTTTGATGGGACCAGACTCCAAACACTGCCCGCTCAGAGAGAGAGCTCCAA 1344  
Oy 329 329  
Db 1345 GTGTTTGAAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGAGACAG 1404  
Oy 329 329  
Db 1405 CTGGCTGACCTCAGCGTCTTCCAGAAACCTGCAGAAATATCCGGGAGAGAAATTTGCACAAT 1464  
Oy 329 329  
Db 1465 GGGCCTACTCGCTGACCTGCAGAGGGCTGGGCAATCAGCTGGGTGGGCTGGCTACATG 1524  
Oy 329 329  
Db 1525 AGGGAATGGGAGTGGAGACTGGCCCTCATCCACATTAACACCCACTGTCTTCTGTGCAC 1584  
Oy 329 329  
Db 1585 ACGTGCCCTGGGACAGCTTTTGGAAACCGGCAACAGCTGTGCTCCACACTGCCAAC 1644  
Oy 329 329

Db 1645 CGGCCAGAGAGACAGTGTGTGGGGCAGAGGCTTGCCCTGCACACAGCTGTGGCCCGGAGG 1704  
Oy 329 329  
Db 1705 CACTGTGGGGGTCCAGGGCCCAACCAAGTGTGTCAACTGCAGCACTGCTTGGGGCAG 1764  
Oy 329 329  
Db 1765 GAGTCCGTGGAGAAATGCCAGTACTGCAGGGGCTCCCAAGGAGATATGTGAATCCAGG 1824  
Oy 329 329  
Db 1825 CACTGTGGCGTGCCACCTGAGTGTGACGCCAGAAATGCTCAGTACCTGTTTGA 1884  
Oy 329 329  
Db 1885 CCGGAGGCTGACCAAGTGTGTGGCTGTGCCCACTATTAAGACCTTCCTTGTGCGTGCC 1944  
Oy 329 329  
Db 1945 CGCTGCCCGAGGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2004  
Oy 329 329  
Db 2005 GAGAGGGCGCATGCCAGCTTGCCCATCAACTGCACCACTCCTGTGTGAGACTGGAT 2064  
Oy 329 329  
Db 2065 GACAAGGCTGCCCGCGGAGCAGAGACCAGCCCTGTGACGTGCATGCTCTGCGGTG 2124  
Oy 329 329  
Db 2125 GTTGCAATTCGTGCTGTGTGCTTTGGGGGTGTGCTTTGGGATCCTCATCAAGGACGG 2184  
Oy 329 329  
Db 2185 CAGCAGAAGATCCGGAAGTACAGATGCGGAGACTGTGCAGAAACGAGAGCTGTGAG 2244  
Oy 329 329  
Db 2245 CCGTGCACACTAGCGGAGCCATGCCCAACAGGCGCAGATCCGATCTT3AAAGACAG 2304  
Oy 329 329  
Db 2305 GAGCTGAGGAAGTGAAGTGTCTTGATCTGGCGCTTTGGCAGAGTACAAAGGCATC 2364  
Oy 329 329  
Db 2365 TGGATCCTGATGGGAGATGTGAATAATCCAGTGGCCATCAAGTGTGAGGGAANAAC 2424  
Oy 329 329  
Db 2425 ACATCCCCCAAAAGCCAAACAAGAAATCTTAGACGAAGCATACGTATGCTGTGTGGC 2484  
Oy 329 329  
Db 2485 TCCCATATGTCTCCCGCTTGTGGCATCTGCTGACATCCAGGGTGCAGTGTGACA 2544  
Oy 329 329  
Db 2545 CAGCTTATGCCCTATGCTGCTCTTAGACATGTCCGGAAACCGCGAGAGCTGTGGC 2604  
Oy 329 329  
Db 2605 TCCAGAGACTGCTGAAGTGTGTGTGACAGATTGCCAAGGGAGTGAAGTACCTGGAGAT 2664  
Oy 329 329  
Db 2665 GTGGGCTGTACACAGGAGCTTGGCCGCTGGAACGTGTGTCAAGATCCCAACAT 2724  
Oy 329 329  
Db 2725 GTCAAAATTAAGACTTGGGCTGGCTGGCTGTGACATTTGACAGAGACAGATGACAT 2784

Oy	329	-----	329
Db	2785	GCAGATGGGGGCAAGGTGCCATCAAGTGGATGCGCTGGAGTCCATTCTCCGCGCGCG	2844
Oy	329	-----	329
Db	2845	TTCAACCACAGAGTGAATGTGTGAGTTATGTGTGACTGTGTGGAGCATGACTTTT	2900
Oy	329	-----	329
Db	2905	GGGGCCAAACCTTACGATGGGATCCAGCCGGGAGATCCGATCCCTGTGGAAAAAGGG	2966
Oy	329	-----	329
Db	2965	GAGGGCGTGGCCAGCCCCCATCTGCACATTGATGTCTACATGATCATGGTCAAAATGT	3022
Oy	329	-----	329
Db	3025	TGGATGATTGACTCTGAATGTGGCCAAAGATTCGGGAGTGGTGTCTGAATTTCTCCGC	3088
Oy	329	-----	329
Db	3085	ATGGCCAGGAGCCCCCAGCGCTTTGTGTATCTCAGAAATGAGACTTGGGCCCACT	3144
Oy	329	-----	329
Db	3145	CCCTTGCACAGACACCTTCTACCGCTCACTGTGGAGAGCATGACATGGGGGACGTGGT	3200
Oy	330	-----	330
Db	3205	GATGCTGAGGAGTACTGGTATCCCCAGCAGGGCTTCTTCTGCACAGACCTTGCCCGCGGC	3266
Oy	331	AlaGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAsp	350
Db	3285	GCTGGGGCATGTGTCACACAGGCACCGACGTCATCTACACAGAGTGGCGGTGGGAGC	3322
Oy	331	LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer	370
Db	3335	CTGACACTAAGGGCTGGAGCCCTTGAAAGAGAGGCCCCAGGTCTCACTGGCACCTTC	3386
Oy	371	GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu	390
Db	3385	GAGGGCGTGGCTCCGATGTATTGTGATGTGTGACTGGGAATGGGGGCACCAAGGCGTG	3444
Oy	391	GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGlyAspProThrVal	410
Db	3445	CAAAAGCTCCCAACATGAGCCAGCCCTCTACACGGTACGTAGAGACCCCAACAGTA	3500
Oy	411	ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu	430
Db	3505	CCCCGCCCCGTGACAGTGAAGGTCTGACGTCAGCCGTCGACGCCCAAGCCTGAA	3566
Oy	431	TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro	450
Db	3565	TATGTGAACACAGCATGTTGCGGCCACGCCCTTCCGCCCGGAGAGGGCCCTGTGCT	3622
Oy	451	AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn	470
Db	3625	GCTGCCCACTGCTGTGTGTGCACCTGTGAAAAGGGCCAAAGACTCTCTCCCAAGGAAGAT	3686
Oy	471	GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr	490
Db	3685	GGGGTGGTCAAAAGACGTTTTTGCTTTGGGGGTGGCCGTGAGAAACCCGAGTACTTGACA	3744
Oy	491	ProGlnGlyGlyAlaAlaProGlnProHisAspProProAlaPheSerProAlaPheAsp	510
Db	3745	CCCCAGGAGAGGTGCCCTCTCAGCCCCCAACCTCTCTGCTTCTCAGCCACGCTTCGAC	3800
Oy	511	AsnLeuTyrTyrTrpAspGlnAspProGluArgGlyValaProProSerThrPheLys	530
Db	3805	AACCTATTACTGTGGACAGAGACCCACAGACGGGGGGGCTTCACCAAGCACTTTCAA	3866

OY	531	gltttrprothAlaIGlnAsnProGluTyrLeuGlyLeuAspVal.PrcIALAIAlaIaPro	550
Db	3865	GGGACACTCAGCGCAGACACCAGATCACTGGGTCTGGAGCTGCCA-----GTGTGA	3918
OY	551	AlaArgSerProSerPro	556
Db	3919	ACCAGAGCCCAAGTCGC	3936
RESULT 7			
LOCUS	BD005474	4530 bp	DNA linear, PAT 31-JAN-2002
DEFINITION	Cellular immunogens useful cancer vaccines.		
ACCESSION	BD005474		
VERSION	BD005474.1	GI:18633845	
KEYWORDS	JP 2001501909-A/4.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 4530)		
AUTHORS	Halpern,M.S. and England,J.M.		
TITLE	Cellular immunogens useful cancer vaccines		
JOURNAL	Patent: JP 2001501909-A 4 13-FEB-2001;		
COMMENT	ALLEGHENY UNIVERSITY OF THE HEALTH SCIENCES		
	OS Unidentified		
	PN JP 2001501909-A/4		
	PD 13-FEB-2001		
	PE 13-JAN-1997 JP 1997526124		
	PR 19-JAN-1996 US 60/010262		
	PI MICHAEL,S HALPERN,JAMES M ENGLAND		
	PC A01K63/00,A61K39/00,A61K39/38,A61K48/00,C12N5/00,C12N15/00 CC		
	Strandness: Single;		
	CC Topology: Linear;		
FEATURES			
source	FT	Location/Qualifiers	'/organism='Unidentified'.
	FT Key	1..4530	
	FT source	Location/Qualifiers	1..4530
BASE COUNT	922 a 1382 c 1346 g 880 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1,49e-80	Length:	4530
Score:	2423.00	Matches:	524
Percent Similarity:	41.14%	Conservative:	5
Best Local Similarity:	40.75%	Mismatches:	21
Query Match:	64.80%	Indels:	736
DB:	6	Gaps:	4
US-09-821-883-2 (1-690) x BD005474 (1-4530)			
OY	2	ARGAIAAlaProLeuLeuLeuAlaArgAIA-----AlaSerLeuSerLeuGly	17
Db	94	CAGCCCCCTCCACAGCGGGGTCCAGCGAGACCATGGGGCCGAGCCGAGACGACATG	153
OY	18	PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGlyLeuAlaArg	37
Db	154	GAGCTGGCGGGCTTGTCGCCGTG-----GGCTCTCCCTGCCCCTGTCGCCCC	204
OY	38	GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer	57
Db	205	GGAGCGCGAGACCCAGCAAGTGTGCACCGCACAGACATGAAAGCTGCGGCTCCTCCAGT	264
OY	58	ProGluTrpHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly	77
Db	265	CCCGAGACCCACCTTGGACATGCTCCGCCACCTTACCAAGGCTGCCAGGTGTGACGGA	324
OY	78	AsnLeuGluLeuThrTyrLeuProThraAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
Db	325	AACCTGGAACCTACCTACTGCGCCACCAATGACACCTGTCTTCTTCGACAGATTTCAG	384
OY	98	GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117

|||||  
Db 385 GAGGTGACGAGGCTACGTCTACCTCCACACCAAGTAGAGAGGTGCCACTGCAGAG 444  
Oy 118 LeuAlaIleValArgIleThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137  
Db 445 CTGGCGATTGTGCGAGGACCCAGCTCTTTGAGGACAACTAAGCCCTGGCTGTAGAC 504  
Oy 138 AsnGluAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157  
Db 505 AATGAGACCCGCTGACAACTACCAACCCCTGTGCACAGGGGCTCCCGCAGAGGCTGGCG 564  
Oy 158 GluLeuGlnLeuArgSerLeuThrGlnIleLeuGlyGlyValLeuIleGlnArgAsn 177  
Db 565 GAGCTGCAGCTTCGAAGCCTCACAGAGATCTGAAGAGGGGCTGTGATCCAGCGGAAC 624  
Oy 178 ProGlnLeuGlySerGlyAspThrIleLeuThrPylAspPheHisIleAsnAsnGln 197  
Db 625 CCGCAGCTCTGCTACAGAGACAGATTGTGGAAGGACATCTTCCACAAGAAACACAG 684  
Oy 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217  
Db 685 CTGGCTTCACACTGATAGACACCAACGCTCTGGGCTGCACCCCTGTTCTCGAATG 744  
Oy 218 CysIlySerArgCysTrpGlyLeuSerGluAspCysGlnSerLeuThrArgThr 237  
Db 745 TGTAGGGCTCCCGCTGCTGGGAGAGAGATTCTGAGAGATTCTCAGAGCCTGACGCGCACT 804  
Oy 238 ValCysAlaGlyIleCysAlaArgCysIlySerGlyProLeuProThrAspCysGlyHisGlu 257  
Db 805 GCTCTGGCGGGTGGCTGTGCCCCGTGCAGAGGGGCCACTGCCACTGACGTGCTCCATGAG 864  
Oy 258 GlnCysAlaAlaGlyCysThrGlyProIlyHisSerAspCysLeuAlaCysLeuHisPhe 277  
Db 865 CAGTGTGGTGGCGGGTGGACGGGGCCCAAGCACTGTGACTGCTGGCTGCCCTCCACTTC 924  
Oy 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297  
Db 925 AACCAAGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTCACTACAAACACAGACAG 984  
Oy 298 PheGluSerMetProAsnProGluArgTyrThrPheGlnAlaSerCysValThrAla 317  
Db 985 TTGAGTGCATGCCCAATCCAGAGGCGGATACATTCGTGGCGCAGCTGTGTGACTGCC 1044  
Oy 318 CysProTyrAsnTyrLeuSerThrAspValGlySer 329  
Db 1045 TGTCCCTACAACTACTCTTCTACGGACGTGGATCCTGCACCCCTGTGCCCCGTGCAC 1104  
Oy 329 329 329  
Db 1105 AACCAAGAGTGCACAGACAGAGATGAAACACAGCGGTGTGAGAGTGCACAAAGCCCTGT 1164  
Oy 329 329 329  
Db 1165 GCCCAGTGTGCTATGTCGTGGGCATGGAGCACTTGGCAGAGGTGAGGGCACTTACCACT 1224  
Oy 329 329 329  
Db 1225 GCCAATATCCAGAGATTGTGCTGGTGCAGAAGATCTTTGGAGGCTGCAATTTCTGCCG 1284  
Oy 329 329 329  
Db 1285 GAGAGCTTTGATGGGACCCAGCTCCAACTGCCCGCTCCAGCCAGACAGAGCTCCAA 1344  
Oy 329 329 329  
Db 1345 GTGTTTGAAGACTCTGTGAAGAGATCAACAGGTACTATACATCTCAGCATGGCCGACACGC 1404  
Oy 329 329 329  
Db 1405 CTGGCTGACCTCAGGCTTCTCAGAACTGCAAGTAATCCGGGAGAGAAATTCTGCACAAAT 1464  
Oy 329 329 329

Db 1465 GCGGCTACTGCTGACCTGACCCCTGCAGAGGCTGGGCATCAGCTGGCTGGGCTGAGCTCACTG 1524  
Oy 329 329 329  
Db 1525 AGGGAATGGGACATGAGTGAAGTGGCCCTCATCCACATTAACCCACCTCTGCTTGTGCAC 1584  
Oy 329 329 329  
Db 1585 ACGGTGCCCTGGGACACAGCTTTTGGAAACCCGACCAAGCTCTCTCCACACTGCACAC 1644  
Oy 329 329 329  
Db 1645 CGGCCAGAGACAGTGTGTGGCGAGGGGCTTGCCCTGCACACACTGTGCGCCCGAGGG 1704  
Oy 329 329 329  
Db 1705 CACTCTGGGGTCCAGGGGCCACCCAGTGTGTCAAACTGACAGCACTCTCTCGGGGCCAG 1764  
Oy 329 329 329  
Db 1765 GAGTGCATGAGAGAAATGCCAGTACTGCAGAGGGGCTCCCGAGGAGTATGTGAATGCACAG 1824  
Oy 329 329 329  
Db 1825 CACTGTTTGCCTGCCACCTGAGTGTGACGCCAGAAATGCTCAGTACCTGTTTGA 1884  
Oy 329 329 329  
Db 1885 CCGAGGCTGACCACTGTGTGCTGTGCCCACTATTAAGAACCTTCCTTCTGCGTGGCC 1944  
Oy 329 329 329  
Db 1945 CGCTGCCCGAGCGGTGTGAACCTGACCTCTCTCATATGCCCATGTGAAATTTCCAGAT 2004  
Oy 329 329 329  
Db 2005 GAGGAGGGCGCATGCCAGCCTTGCCCATCAACTGACACCACCTCTGTGTGAGACTGTGAT 2064  
Oy 329 329 329  
Db 2065 GACAAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTGTGACTGCTCTCTGCGGTG 2124  
Oy 329 329 329  
Db 2125 GTTGCATTTCTGCTGTGTGCTTGGGGGTGCTTTGGGATCCTCATCAAGCAGCG 2184  
Oy 329 329 329  
Db 2185 CAGCAGAAAGATCCGGAAGTACACGATGCCGAGACTGTGCAGAAACGAGAGCTGTGGAG 2244  
Oy 329 329 329  
Db 2245 CCGCTGACACCTAAGCGGAGCATGCCCAACAGGGCGAGATCGGATCCTGAAAGAGAG 2304  
Oy 329 329 329  
Db 2305 GAGCTGAGAAAGTGAAGTGTGATCTGCGCTTTTGGCAGATCTACAAG36CATC 2364  
Oy 329 329 329  
Db 2365 TGGATCCTGTAGGGAGAAATGTGAATTTCCATGGGCCATCAAGTGTGAGGAGAAAC 2424  
Oy 329 329 329  
Db 2425 ACATCCCCCAAGCCAAACAAAGAAATCTTAGACGAAGCATACGTATGCTGTGTGGGC 2484  
Oy 329 329 329  
Db 2485 TTCCCATATGTCTCCGCTTGTGGCATCTGCTGACATCCAGAGTGCATTTGGTGAACA 2544  
Oy 329 329 329  
Db 2545 CAGCTTATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAACCGGAGCGCTGGGC 2604

QY 329 ----- 329  
 Db 2605 TCCAGAGACCTGCTGAAGTGTATGACATTCGCAAGGGGATGAGTACTGAGAGAT 2604  
 QY 329 ----- 329  
 Db 2665 GTGGGCTCGTACACAGGAGTGTGGCGCTGGAAAGCTGTGTCAGAGTCCCAACCAT 2724  
 QY 329 ----- 329  
 Db 2725 GTCAAAATTACAGACTTCGGGCTGGCTGGCTGTCGACATTGACAGACAGAGTACCAT 2784  
 QY 329 ----- 329  
 Db 2785 GCAGATGGGGCAAGTGGCCCATCAAGTGAATGGCGCTGGAGTCCATTCTCCGGCGG 2844  
 QY 329 ----- 329  
 Db 2845 TTCACCCACAGAGTGTGTGTGAGATTATGTGTGACTGTGTGGAGCTGATGACTTTT 2904  
 QY 329 ----- 329  
 Db 2905 GGGGCAAACTTACGATGGATGCCAGCCGGAGATCCCTGACTGCTGGAAGAGGG 2964  
 QY 329 ----- 329  
 Db 2965 GAGGGCTGCCCCAGCCCCCATCTGCACATGATGTCTACATGATCATGTCAAAATGT 3024  
 QY 329 ----- 329  
 Db 3025 TGATGATGACTGTGAATGTGCGCAAGATTCCGGAGTGTGTGTAATTCCTCCGC 3084  
 QY 329 ----- 329  
 Db 3085 ATGGCCAGGAGCCCCAGCGCTTGTGTGTCATCCAGATAGAGACTTGGCCAGCCACT 3144  
 QY 329 ----- 329  
 Db 3145 CCCTTGGACAGACCTTCTACCGCTCACTGCTGAGAGCATGATGAGGGGAGCTGTG 3204  
 QY 330 ----- Gly 330  
 Db 3305 GATCTGAGAGATGTCTGGTACCCAGAGGGCTTCTTCTGTCCAGACCTCCCGCCG 3264  
 QY 331 AlaGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAsp 350  
 Db 3365 GCTGGGGCATGGTCCACACAGGACCGGACGCTCATCTACAGAGTGGCGGTGGGAC 3324  
 QY 351 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 370  
 Db 3325 CTGACACTAGGGCTGAGAGCCCTCTGAAGAGAGGCCCCAGGCTCCACTGGCACCTCC 3384  
 QY 371 GluIleValIleGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeu 390  
 Db 3385 GAAAGGGCTGGCTCGAATGATTTGATGTGACTGGGAATGGGGAGCCAAAGGGCTG 3444  
 QY 391 GluSerLeuProThrHisAspProSerProLeuGluAlaArgTyrSerGluAspProThrVal 410  
 Db 3445 CAAAGCCTCCCAACATGACCCAGCCCTCTACAGCGGTACATGAGAGACCCACAGTA 3504  
 QY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430  
 Db 3505 CCCCTGCCCTCTGAGACTGATGCTGAGCTGAGTGGCCCCCTGACCTGACAGCCCAAGCTGAA 3564  
 QY 431 TyrValAsnGlnProAspValArgProGluProProSerProAlaGlyGlyProLeuPro 450  
 Db 3565 TATGTGAACACGACGAGATGTTGGCCCCAGCCCCCTTGGCCCCAGAGGGCCCTCTCCCT 3624  
 QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaIleThrLeuSerProGlyIleAsn 470  
 Db 3625 GCTGCCGACCTGCTGTGGTGCACACTCTGTGAAAGGCCAAAGACTCTCTCCCAAGGAGAAAT 3684

QY 471 GlyValIleValIleAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490  
 Db 3685 GGGTCTGCAAAAGACGTTTTCCTTTGGGGGTCCCGGAGAACCCCGAGTACTTGACA 3744  
 QY 491 ProGlnIleGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp 510  
 Db 3745 CCCAGGAGAGAGCTGCCCTCCACGCCCTCTCTGCTTGCACGCCAGCTTGGAC 3804  
 QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530  
 Db 3805 AACCTTATTACTGGACACAGACCCACAGAGCGGGGCTCCACCCAGCAGCTTCAAA 3864  
 QY 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaPro 550  
 Db 3865 GGGACACTACGGCAGAGAACCCAGAGTACTGTGCTGAGACGTGCA-----GTGTGA 3918  
 QY 551 AlaArgSerProSerPro 556  
 Db 3919 ACCAGAGGCCCAAGTCCG 3936  
 RESULT 8  
 LOCUS 121124 4530 bp DNA linear PAT 07-Oct-1996  
 DEFINITION Sequence 9 from patent US 551885.  
 ACCESSION 121124  
 VERSION 121124.1 GI:1601478  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 4530)  
 AUTHORS Razindin and Sarkar, F.H.  
 TITLE ERB2 promoter binding protein in neoplastic disease,  
 JOURNAL Patent: US 551885-A 9 21-MAY-1996;  
 FEATURES  
 source 1..4530  
 location/Qualifiers  
 BASE COUNT 922 a 1382 c 1346 g 880 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.49e-80 Length: 4530  
 Score: 2423.00 Matches: 524  
 Percent Similarity: 41.14% Conservative: 5  
 Best Local Similarity: 40.75% Mismatches: 21  
 Query Match: 64.80% Indels: 736  
 Gaps: 4  
 US-09-821-883-2 (1-690) x 121124 (1-4530)  
 QY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17  
 Db 94 CGGCGCTCCACAGCGGGGTCCAGCGGAGCATGGGCGCGAGCGGAGGAGACCATG 153  
 QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37  
 Db 154 GAGCTGGCGGCTGTGCTCCGCTG-----GGCTCTCTCTCGCCCTCTTGGCCCCC 204  
 QY 38 GlyValIleAspSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57  
 Db 205 GGAGCCCGAGACACCCCAAGTGTGCACCGGACACACATGAAAGTCCGCGCTCCCTCCAGT 264  
 QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValIleGlnGly 77  
 Db 265 CCGAGAGACCACTCGACATAGTCTCGGCACTTACAGAGGCTCCAGGTGTGCAGAGA 324  
 QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97  
 Db 325 AACCTGGAACCTACACTACCTGCGCCACCAATGCCAGCTGTCTTCGTCAGAGATATCCAG 384  
 QY 98 GluValGlnGlyTyrValIleValIleHisAsnGlnValArgGlnValProLeuGlnArg 117  
 Db 385 GAGGTGAGGGCTACGTGCTCATCTGCACAAACCAATGAGGAGGTGCCACATGCAAGAG 444

OY	118	LeuArgIleValaArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp	137
Db	445	CTGGGAGATTGGCCAGGACCACCAAGCTCTTTGAGAGCAACATATGAGCCCTGGCGGTCTATAC	504
OY	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyValLeuArg	157
Db	505	AATGAGAGCCCGGTGAACCATATACCACCCCTGTCACAGGGGGCTCCCCAGAGAGCCGTGGG	564
OY	158	GluLeuGlnLeuAsnArgSerLeuThrGluIleLeuValGlyValLeuIleGlnArgAsn	177
Db	565	GAGGTGACGCTTCACAAACCCACAGAGATCTTGAAAGGAGGGGTCTGTATCCAGGGAGAC	624
OY	178	ProGlnLeuCystrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGln	197
Db	625	CCCCAGCTCTGCTACAGAGACAGCATTTTGTGAAAGACATCTTCCACAAAGAACACAG	684
OY	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProwet	217
Db	685	CTGGCTCTCACACTGATAGACACCAACCCGCTCTGGGGCTGGCACCCCTGTTCTCCGATG	744
OY	218	CysLysGlySerArgCysTyrPglYgluSerSerGluAspCysGlnSerLeuThrArgThr	237
Db	745	TGTAAAGGGCTCCCGCTCTGGGAGAGAGTCTGAGAGATTGTCAAGAGCTTGACGGGACT	804
OY	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysGlyHisGlu	257
Db	805	GTCGTGGCGGTGGCTGTGGCCGCTGGAAGGGGCCACTGCCACTGACGTGCTGCCATAG	864
OY	258	GlnCysAlaAlaArgCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe	277
Db	865	CAGTGTGCTGGCGGCTCACAGGGGCCCAACACACTGTGACTGGCTGGCCCTCCACTTC	924
OY	278	AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297
Db	925	AACCAACAGTGGCATCTGTGACTGTGCATGCCACGCCCTGGTCACTACACACACAGACG	984
OY	298	PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla	317
Db	985	TTTGTGTCATGACCCCAATCCCGAGGGCGGTATATACATTTGGGGCCGACTGTGTACTGCC	1044
OY	318	CysProTyrAsnTyrLeuSerThrAspValGlySer-----	329
Db	1045	TGTCCCTACAACTACTCTTCTACGAGAGTGGATCTTGACACCCTGTGCCCCCTGCAC	1104
OY	329	-----	329
Db	1105	AACCAAGAGTGAACAGAGAGATGGAACACACGGGTGTGAGAAAGTGCACGAAGCCGTG	1164
OY	329	-----	329
Db	1165	GCCCGAGTGTCTATGTCTTGGGCATGAGACACTTGCAGAGGTGAGGCACTTTACAGT	1224
OY	329	-----	329
Db	1225	GCCCAATATCCAGAGATTGCTGGCTGCAGAGAGATCTTTGGAGACCTGGCATTTCTGCCG	1284
OY	329	-----	329
Db	1285	GAGAGCTTTGATGGGAGACCAGCCGCTCCACACTGTGCCCGCTCCAGCAGAGAGCTCCA	1344
OY	329	-----	329
Db	1345	GTTGTTGAGACTCTGGAAGAGATDACAAGTTACTCTTATCATCTCAAGCATGGCCGAGACG	1404
OY	329	-----	329
Db	1405	CTGCGCTGACCTCAGACGTCTTCCAGAAACCTGCAAGTATATCCGGGAGAGATTTGTGACAA	1464
OY	329	-----	329
Db	1465	GAGCGCTACTGCTGACCTTGCAAGGGGCTGGGAGTCAAGTGGCTGGGGCTGCGCTCACTG	1524

0Y	329	-----	329	-----	329
Db	1525	AGGGAACGGGCACTGGACATGGCCCTCATTCACACATAAACCCACCTCTGCTGTGCAC	1584		
0Y	329	-----	329	-----	329
Db	1585	ACGGTGCCTGGGACACGCTCTTTGGGAACCGGACCAAGCTCTGCTCCACACTGCCAAC	1644		
0Y	329	-----	329	-----	329
Db	1645	CGGCCAGAGACGAGTGTGTGGCGAGGGCCTGGCCTGCCACACAGCTGTGGCCGAGGG	1704		
0Y	329	-----	329	-----	329
Db	1705	CACGTCTGGGGTCCAGGGCCACCACAGTGTTCACCTGCAGCCAGTTCTTCGAGGCCAG	1764		
0Y	329	-----	329	-----	329
Db	1765	GAGTCCGTGGAGAAATCCGAGTACTGACAGGGGCTCCGAGGAGTATGTGAATGCCAGG	1824		
0Y	329	-----	329	-----	329
Db	1825	CACGTTTGGCCGTGCACACCTGAGTGTCAAGCCCCAGAAATGGCTCAGTGACCTGTTTGA	1884		
0Y	329	-----	329	-----	329
Db	1885	CCGGAGCTGACACAGTGTGTGGCCTGTGCCACTATAGAGCCCTCCCTTGTGCTGGCC	1944		
0Y	329	-----	329	-----	329
Db	1945	CGCTGCCCAAGCGGTGTGAACCTGACCTTCCTACATGCCCATCTGGAATTTCCAGAT	2004		
0Y	329	-----	329	-----	329
Db	2005	GAGGAGGGCGATGCGACGCTTGCCCATCAACTGCACCCACTCCGTGTGTGACCTGGAT	2064		
0Y	329	-----	329	-----	329
Db	2065	GACAAGGGCTGCCCGCGGACGAGAGAGCCAGCCCTGTGACCTCATGCTCTTCCGGTG	2124		
0Y	329	-----	329	-----	329
Db	2125	GTTGGCATTCTGTGTCGTGTGTTGGGGGTGTTTGGAGTCTCATCAAGCGCGG	2184		
0Y	329	-----	329	-----	329
Db	2185	CAGCAGAAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGAAAAGGAGCTGGTGAG	2244		
0Y	329	-----	329	-----	329
Db	2245	CCGCTGACACCTAGCGGAGGATGCCCAACAGGCGCAGATGCGGATCCTGAAGAAGAG	2304		
0Y	329	-----	329	-----	329
Db	2305	GAGCTGAGAGAGTGAAAGGTGCTTGGATCTGGCGCTTTTGGCAGAGTCTCAAGGGCATC	2364		
0Y	329	-----	329	-----	329
Db	2365	TGGATCCCTGATGGGGGGAATGTGAATAATCCAGTGGCCATCAAAAGTGTGAAGGAAC	2424		
0Y	329	-----	329	-----	329
Db	2425	ACATCCCCCAAGCCAAACAAGAAATCTTAGAGGAAGCATACGTGATGGCTGTGTGGGC	2484		
0Y	329	-----	329	-----	329
Db	2485	TCCCATATGTCTCCGCGCTTGGGCAATCGCTGACATCCACGGTGCAGCTGTGTGACA	2544		
0Y	329	-----	329	-----	329
Db	2545	CAGCTTATGCCCTATGGCTGCCCTTAAAGACCATGTCCGGAAAAACCGGAGCGCTGGGC	2604		
0Y	329	-----	329	-----	329

Db 2605 TCCAGGACCTGCTGAAGTGTGTATGACAGATTGCCAAGGGGATGACCTACTGAGAGAT 2664  
 Qy 329 ----- 329  
 Db 2665 GTGGGCTCGTACACAGAGACTTGGCGCTCGGAACGTGCTGTCAGAGATCCCAACCAT 2724  
 Qy 329 ----- 329  
 Db 2725 GTCMAAATACAGACTTCGGGCTGGCTCGGCTGTGACATTCAGAGAGAGATACCAT 2784  
 Qy 329 ----- 329  
 Db 2785 GCAGATGGGGCAAGTGGCCATCAAGTGGATGGCTGTGAGTCCATTCGCCGGCGG 2844  
 Qy 329 ----- 329  
 Db 2845 TTCACCCACCAAGATGATGTGTGAGATTATGTGTGACTGTGTGGAGCTGATGACTTTT 2904  
 Qy 329 ----- 329  
 Db 2905 GGGGCCAAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACTGTGAAAGGGG 2964  
 Qy 329 ----- 329  
 Db 2965 GAGCGGCTGCCCCAGCCCCCATCTGCACCATTTGATGCTACATGATGATGTCATAATGT 3024  
 Qy 329 ----- 329  
 Db 3025 TGGATGATGACTCTGAATGTGGCCCAAGATTCGGGAGTTGGTGTGAATTCCTCCGC 3084  
 Qy 329 ----- 329  
 Db 3085 ATGGCCAGGAGACCCAGCGCTTGTGTGATCATCAGATGAGACTTGGGCCAGCCAGT 3144  
 Qy 329 ----- 329  
 Db 3145 CCCTTGACAGACACTTCTACCGCTCAGCTGTGAGAGACATGATGGGGAGCTGTG 3204  
 Qy 330 ----- Gly 330  
 Db 3205 GATGCTGAGAGATCTGTGTACCCAGCAGAGGCTTCTTGTCCAGACCTGCCCGGGC 3264  
 Qy 331 AAGGlylymetvalhshisatghisatgsersestheratgserglylyasp 350  
 Db 3265 GCTGGGGCATGTGCCACAGAGCAGCGCATCTACAGAGATGGCGGGGAGC 3324  
 Qy 3351 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 370  
 Db 3325 CTGACACTAGGGCTGGAGCCCTCTGAAGAGAGAGCCCCCAGGTCTCCACTGGCACCTCC 3384  
 Qy 3371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeu 390  
 Db 3385 GAAGGGGCTGGCTCCGATGTATTGATGTGACTGGGAATGGGGAGCAGCAAGGGGCTG 3444  
 Qy 3391 GluSerLeuProThrHisAspProSerProLeuGluAlaArgTyrSerGluAspProThrVal 410  
 Db 3445 CAAGGCTCTCCACACATGAGCCCGCTCTACAGGGGATGAGAGAGCCACAGTA 3504  
 Qy 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430  
 Db 3505 CCCCTGCCCTGTAGATGATGGCTAGCTGCTCCCTGACTGACGCCCCCAGCCCTGAA 3564  
 Qy 431 TyrValAsnGluProAspValArgProGluProProSerProArgGluGlyProLeuPro 450  
 Db 3565 TATGTGAACAGCAGATGTGGGCCAGCCCTCGCCGAGAGAGGGCCCTGCTCCT 3624  
 Qy 451 AlaAlaArgProAlaGlyAlaThrLeuGluAlaAlaGlyThrLeuSerProGlyLysAsn 470  
 Db 3625 GCTGGCGGACCTGTGTGCTGAGTGAAGGGCCCAAGACTCTCTCCAGGAGAAAGAT 3684  
 Qy 471 GLVAlaValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490  
 |||||||

Db 3685 GGGGCTGCTCAAGACGTTTTTGCCTTTGGGGGTCCTGTGAGAGACCCGAGTACTTGACA 3744  
 Qy 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510  
 Db 3745 CCCAGGGAGAGAGTGGCCCTCAGCCCGCCACTCTCTCTGCTTCCACCCAGCCTTGAC 3804  
 Qy 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530  
 Db 3805 AACCTTATTACTGAGACAGGAGACCCAGAGAGCGGGGCTCCACCCAGCCTTCAAA 3864  
 Qy 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaPro 550  
 Db 3865 GGGACACTTACGGCAGAGAACCCAGAGTACCTGGCTGTGAGCTGCCA-----GTGTGA 3918  
 Qy 551 AlaArgSerProSerPro 556  
 Db 3919 ACCAGAGGCCAAGTCCG 3936  
 RESULT 9  
 159745  
 LOCUS 159745 4530 bp DNA linear PAT 07-OCT-1997  
 DEFINITION Sequence 9 from patent US 5654406.  
 ACCESSION 159745  
 VERSION 159745.1 GI:2478377  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 4530)  
 AUTHORS Razuddin and Sarkar, F. Hogue.  
 TITLE Antibody to ERBB2 promoter binding factor  
 JOURNAL Patent: US 5654406-A 9 05-AUG-1997;  
 FEATURES  
 source 1..4530  
 BASE COUNT 922 a 1382 c 1346 g 880 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,49e-80 Length: 4530  
 Score: 2423.00 Matches: 524  
 Percent Similarity: 41.14% Conservative: 5  
 Best Local Similarity: 40.75% Mismatches: 21  
 Query Match: 64,808 Indels: 736  
 DB: Gaps: 4  
 US-09-821-883-2 (1-690) x 159745 (1-4530)  
 Qy 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17.  
 Db 94 CGGCGCTCTCCAGCGGGGCTCCAGCGGAGCATGGGCGCGAGCGCGAGTGAAGCACCATG 153  
 Qy 18 PheLeuPheLeuLeuPhePheTyrPheAspArgSerValLeuAlaLysGluLeuAlaArg 37  
 Db 154 GAGCTGGCGGCTTGTCTCCCTG-----GGCTCTCTCTCGGCCCTCTTGGCCGCC 204  
 Qy 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57  
 Db 205 GGAGCCCGGAGCACCCCAAGGTGTGACCGGACAGACATGAAGGTGCGGCTCCCTGCCAGT 264  
 Qy 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValAlaGlnGly 77  
 Db 265 CCCGAGACCCACCTGGACATGCTCCGCGCACCTTACACAGGGCTGGCAGGTGTCAGGGA 324  
 Qy 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97  
 Db 325 AACCTGGAATCTACATCTGCTGCGCCACCAATGCCAGCTGTCTTCTGCGAGATATCCAG 384  
 Qy 98 GluValGlnGlyTyrValLeuLeuAlaHisAsnGlnValAlaArgGlnValProLeuGlnArg 117  
 Db 385 GAGGTGACAGGCTACGTCATCGCTCAGACACCAAGTGAAGGAGGTGCCACATGCAAGAG 444  
 Qy 118 LeuArgIleValAlaArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137  
 |||||||

|||||  
Db 445 CTGCGGATTTGCGAGGACCCAGCCTTTGGAGCAACATGATGCGCCCTGCTAGAC 504  
OY 138 AsnGlyAspProIleuAsnThrThrProVal1ThrGlyAlaSerProGlyGlyLeuArg 157  
Db 505 AATGAGAACCCGCTAACCAATACCAACCCCTGTGCACAGGGGCTCCCGAGAGCCCTGGCG 564  
OY 158 GluLeuGlnLeuArgSerLeuThrGluLeuLysGlyValLeuIleGlnArgAsn 177  
Db 565 GAGCTGACAGCTTCGAAAGCTCCACAGATCTTGAAAGAGAGGGCTTTGATCCAGGGAAAC 624  
OY 178 ProGlnLeuCysTyrGlnAspThrIleLeuTrrLysAspIlePheHisLysAsnGln 197  
Db 625 CCCCAGCTCTGCTACACAGACACGATTTTGTGGAGGACATCTTCCACAAAGAACACAG 684  
OY 198 LeuAlaLeuThrIleuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217  
Db 665 CTGGCTTCACACTGATGACACCAACCGCTCTCGGCCCTGCACCCCTGTCTCCGATG 744  
OY 218 CysLysGlySerArgCysTrrPglLysLeuSerGluAspCysGlnSerLeuThrArgThr 237  
Db 745 TGTAAAGGCTCCCGCTGCTGGGGAAGAGTTGTAGGATTTGTCAAGCCTGACGCGCACT 804  
OY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu 257  
Db 805 GTCTGTCCGCGTGTGCTGCCGCTGCAAGGGGCACTGCCACTGACTGTGCTGCATGAG 864  
OY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277  
Db 865 CAGTGTGTGCGCGCTGACAGGGGCCCCAAGCCTGTGACTGCTGCTGCTCCACATTC 924  
OY 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuVal1ThrTyrAsnThrAspThr 297  
Db 925 AACCAAGATGGATGTGTAGCTGACCTGCCAGCCCTGTGTAACCTACACACACAGAG 984  
OY 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysVal1ThrAla 317  
Db 985 TTTGAGTCCATGCCCAATCCCGAGGGCGGTATACATTGCGGCCAGGTGTGTGATGCC 1044  
OY 318 CysProTyrAsnTyrLeuSerThrAspVal1GlySer 329  
Db 1045 TGTCCCTACAACTACTTTCTAGGAGCTGGGATCTGCAACCCTGTCTGCCCTGCAC 1104  
OY 329 329  
Db 1105 AACCAAGAGGTGACAGAGAGATGAAACACAGCGGTGTAGAGATGCCAAGACCTGT 1164  
OY 329 329  
Db 1165 GCCCGAGTGTCTATGTCTGGGCATGTGAGCACTTGCAGAGGTGAGGGCACTTACAGT 1224  
OY 329 329  
Db 1225 GCCAATATCCAGAGATTGTGCTGCTGCAGAAAGATCTTTGGAGGCTTGCAATTTCTGCCG 1284  
OY 329 329  
Db 1285 GAGAGCTTTGATGGGGAACCCACCTCCAAACACTGCCCGCTCCAGCCAGAGAGCTCCAA 1344  
OY 329 329  
Db 1345 GTGTTTGAGACTCTGGAAGAGATCACAGTTTACTATACATCTCAGCATGCGCGGAGAC 1404  
OY 329 329  
Db 1405 CTGCTGACCTTCAGCGCTTTCAGAACTTCGAATATCCGGGGAGCAATTTCTGCACAA 1464  
OY 329 329  
Db 1465 GCGCGCTACTGCTGACCTGCAAGGGGCTGGGCATCAGTGTGCGGTGGCGTGCACAG 1524  
OY 329 329

Db 1525 AGGAACCTGGCGAGTGAAGTACGTGGCCCTCATTCACACATTAACACCCACCTGTCTGTGCAC 1584  
OY 329 329  
Db 1585 ACGTGCCCTGGGACCACTCTTTTGGAAACCCGACACCAAGCTTGTCTCAGACTTSCAAC 1644  
OY 329 329  
Db 1645 CGGCCAGAGACAGTGTGTGGCGAGGGCCGTGGCTGCACACACCTGTGGGCCGAGAG 1704  
OY 329 329  
Db 1705 CACTGTGGGGTCCAGGGCCACCCAGTGTGTCAACTGACAGCAGTCTTCTGGGGCAG 1764  
OY 329 329  
Db 1765 GAGTGTGAGAGAAATGCCAGTACTGACAGAGGGCTCCCAAGGAGATGTCAATGCCAG 1824  
OY 329 329  
Db 1825 CACTGTTTGCCTGCCACCTGAGTGTACAGCCCAAGATGGCTCAGTACCTGTTTGA 1884  
OY 329 329  
Db 1885 CCGAGGCTGACAGTGTGTGCTGTGCCACTATTAAGAACCTCCCTTGTGCTGGCC 1944  
OY 329 329  
Db 1945 CGCTGCCCAAGCGGTGTGAACCTGACCTTCTCTACATGCCANTGTGAAGTTTCCAGAT 2004  
OY 329 329  
Db 2005 GAGAGGGGCGATGCCACGCTTGCCCATCAACTGCACCCACTCTGTGTGAGACTGGAT 2064  
OY 329 329  
Db 2065 GACAAGGCTGCCCGCCGAGCAGAGACCAAGCCCTGTAGCTCATGCTGTGCGGTG 2124  
OY 329 329  
Db 2125 GTTGGCATTTGCTGTGTGTGTGCTTTGGGGGTGTCTTTGGGATCTCATCAAGGAGCG 2184  
OY 329 329  
Db 2185 CAGCAGAAGATCCGGAATACACGATGCCGAGACTGTGCAGAAACGAGCTGTGGAG 2244  
OY 329 329  
Db 2245 CCGCTGACACTAGCGGAGCGATGCCAACCAAGCGCAGATGCCGATCCTGAAGAGACG 2304  
OY 329 329  
Db 2305 GAGCTGAGGAAGTGAAGTGTGCTTGATCTGGCGCTTTTGGCACAGTCTACAAAGGCATC 2364  
OY 329 329  
Db 2365 TGGATCCCTGATGGGAGAAATGTGAATAATCCAGTGGCCATCAAGTGTGAGGAAAC 2424  
OY 329 329  
Db 2425 ACATCCCCCAAGCCACAAAGAAATCTTAGACGAAGATAGTGTGCTGTGGGC 2484  
OY 329 329  
Db 2485 TCCCATATGTCTCCCGCTTCTGGGCATCTGCTGACATCCAGGCTGACAGTGTGACA 2544  
OY 329 329  
Db 2545 CAGTTATGCTTATGCTGCTTATGACCATGTCCGGAACCGCGAGCCTGGGC 2604  
OY 329 329  
Db 2605 TCCAGAGACTGTGACTGATGTATGAGATTGCCAAGGGAGTACCTTGAGAGAT 2664







Db 2005 GAGAGGGCGCATGCCAGCCTTGCCCATCACTGCACCCACTTCCTGTGTGGACTGGAT 2064  
QY 329 ----- 329  
Db 2065 GACAAAGGCTGCCCGCCGAGACAGAGACCCTCTGACGTCATCGTCTGTGGGTG 2124  
QY 329 ----- 329  
Db 2125 GTTGGCATTCCTGCTGTGCTGTGCTTGGGGGTGGTCTTTGGGATCTCATCAAGCGAGG 2184  
QY 329 ----- 329  
Db 2185 CAGCAGAAAGATCCGGGAAGTACACGATGCGGAGACTGCTGCAGAAACGAGCTGGTGAG 2244  
QY 329 ----- 329  
Db 2245 CCGCTGACACTTACGGGAGGGATGCCAACACGAGCGAGATGGGATCTGAAAGAGAGG 2304  
QY 329 ----- 329  
Db 2305 GAGCTGAGAAAGGTGAAGGTGCTTGGATCTGCGGCTTTTGGCAGTCTACAAAGGCGATC 2364  
QY 329 ----- 329  
Db 2365 TGGATCCCTGATGGGGAGAAATGTGAATAATTCAGTGGCCATCAAAAGTGTGGAGGAAAC 2424  
QY 329 ----- 329  
Db 2425 ACATCCCCCAAGGCAACAAAGAAATCTTAGACGAACATACGATGGTGGTGTGGG 2484  
QY 329 ----- 329  
Db 2485 TCCCATATGTCTCCGCCCTTCTGGGCATCTGCTGACATCAAGGTGACGTGGTGACA 2544  
QY 329 ----- 329  
Db 2545 CAGCTTATGCCCTATGGCTGCTCTTAGACATGTCCGGGAAACCGCGACGCTGGGC 2604  
QY 329 ----- 329  
Db 2605 TCCAGAGACTGCTGAATGATGTGATGACAGATTCGCCAAGGGGATGACTACGTGGAGAT 2664  
QY 329 ----- 329  
Db 2665 GTGGGGCTCCTACACAGGAGCTGGCCGCTCGGAACGTGTGTCAGAGTCCCAACCAT 2724  
QY 329 ----- 329  
Db 2725 GTCAAAATTAACAGACTTCGGGGTGGCTGCTGCTGACATTGACGAGACAGATACCAT 2784  
QY 329 ----- 329  
Db 2785 GCAGATGGGGGCAAGGTGCCCATCAAGTGGATGCGCTGAGTCCATCTCCGCGCGG 2844  
QY 329 ----- 329  
Db 2845 TTACCCACCACAGAGTATGTGTGAGATTATGTGTGACTGTGTGGAGCTGATGACTTTT 2904  
QY 329 ----- 329  
Db 2905 GGGGCCAAACCTTACGATGGGATCCCAAGCCGGGAGATCCCTGACCTGTGGAAGAGGG 2964  
QY 329 ----- 329  
Db 2965 GAGCGGTGCCCGACGCCCATCTGCACCATGTATGTACATGATCATGTCAATGT 3024  
QY 329 ----- 329  
Db 3025 TGGATGATTGACTCTGTGANTGCGCCAAAGTTCCGGAGTTGTGTCTGAATTCCTCCGC 3084  
QY 329 ----- 329  
Db 3085 ATGGCAGGAGCCCGCAGCGCTTTGTGTGATCCAGAAATGAGACTTTGGGCCACGCACT 3144

QY 329 ----- 329  
Db 3145 CCCTTGGACAGACCTTCTACCCGCTCACTGTGGAGAGATGATGAGGGGACTGGTG 3204  
QY 330 ----- 330  
Db 3305 GATGCTGAGGAGATATCTGTAACCCACAGAGGCTTCTTCTGTCCAGACCTGCCCGCGG 3264  
QY 331 AlaGlyIymeValHSHLSArgHLSArgSerSerThrArgSerGlyGlyAsp 350  
|||||  
Db 3265 GCTGGGGCATGTGCCACACAGGACACCGACGATCATCTACAGAGATGGCGGGGAG 3324  
QY 351 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 370  
|||||  
Db 3325 CTGACACTAGAGGCTGAGCCCTGGAAGAGAGGCCCCAGGTCTCCACTGACACCTCC 3384  
QY 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyAlaAlaGlyLeu 390  
|||||  
Db 3385 GAAGGGCTGGCTCCGATGATTTGATGTGACTGGGAATGGGGGAGCCAAAGGGCTG 3444  
QY 391 GlnSerLeuProThrHLSArgProSerProLeuGlnArgTySerGlyAspProThrVal 410  
|||||  
Db 3445 CAAAGCTCCCGACACATGACCCAGCCCTCTACAGCGGTACAGTGAAGACCCACAGTA 3504  
QY 411 ProLeuProSerGluThrAspGlyTyValAlaProLeuThrCysSerProGlu 430  
|||||  
Db 3505 CCCCTGCCCTGTGAGACTGATGGCTAGCTTGGCCCCCTGACCTCAGCCCGCCAGCTGAA 3564  
QY 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 450  
|||||  
Db 3565 TATGTGAACACGACAGATGTTGGCCCCAGCCCTTGGCCCCGAGAGGGGCGCTGTGCT 3624  
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn 470  
|||||  
Db 3625 GCTGCCGAGACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3684  
QY 471 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyLeuThr 490  
|||||  
Db 3685 GGGGTGCTCAAGACGTTTGGCTTGGGGGTGCTGCGAAGACCCGAGTCTTGACA 3744  
QY 491 ProGlnGlyLysAlaAlaProGlnProHLSArgProProProAlaPheSerProAlaPheAsp 510  
|||||  
Db 3745 CCCAGGAGAGAGCTGCCCTCAGCCCCAGCTCTCTCTGCTTCAAGCCAGCTTGAC 3804  
QY 511 AsnLeuTyTrpPaspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530  
|||||  
Db 3805 AACCTATTAATGAGGACACAGAGCCACAGAGGGGGGCTCCACCCAGCAGCTTCAAA 3864  
QY 531 GlyThrProThrAlaGluAsnProGluTyLeuGlyLeuAspValProAlaAlaAlaPro 550  
|||||  
Db 3865 GGGACACCTAGCGCAGGAACCCAGAGTACTGGTCTGAGACGTGCA-----GTGTGA 3918  
QY 551 AlaArgSerProSerPro 556  
|||||  
Db 3919 ACCAGAAGGCCAAGTCCG 3936

RESULT 11  
AX060703 9274 bp DNA linear Pat 22-JAN-2001  
AX060703  
LOCUS Sequence 1 from Patent W00100244.  
DEFINITION  
ACCESSION AX060703  
VERSION AX060703.1 GI:12406100  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequences.  
REFERENCE  
1 (bases 1 to 9274)  
AUTHORS Erickson, S. and Schwall, R.  
TITLES Methods of treatment using anti-erbB antibody-maytansinoid  
conjugates  
JOURNAL Patent: WO 0100244-A.1 04-JAN-2001;  
Genentech, Inc. (US)

FEATURES  
source  
Location/Qualifiers  
1. .9274  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="vector sequence"

BASE COUNT 2152 a 2538 c 2489 g 2095 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,19e-80 Length: 9274  
Score: 2422.00 Matches: 528  
Percent Similarity: 41.12% Conservative: 9  
Best Local Similarity: 40.43% Mismatches: 31  
Query Match: 64.78% Indels: 738  
Gaps: 6

US-09-821-883-2 (1-690) x AX060703 (1-9274)

QY 2 ArgAlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 21  
Db 1695 CGCCCCGGCCCTCCAGCCGGGTCCAGCCGGAGCCATGGAGCTGGCG-----GCC 1745

QY 22 LeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSer 41  
Db 1746 TTGTGCGCGCTGG-----GGGCTCTCTCTGCGCCCTCTGCCCCCGGAGCCGCGAGC 1796

QY 42 ThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHis 61  
Db 1797 .ACCCAAAGTGTACCCGACACAGACATGAAGCTGGCGCTCCCTGCCAGTCCCGAGACCCAC 1856

QY 62 LeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGluLeu 81  
Db 1857 CTGGAACATGTCTCCGACACTCTACACAGGGGTGCCAGGTGGTGCAGAGGAACCTGGAACTC 1916

QY 82 ThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnValGlnGly 101  
Db 1917 ACCTACCTGCCCCCAAGGCCAGCTGCTCTGCGAGATATCCAGAGAGTGCAGGGC 1976

QY 102 TyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIleVal 121  
Db 1977 TACGTGCTCATCGCTCACACACAGTAGAGGAGTCCACAGTCCAGAGGCTCGGATTTTG 2036

QY 122 ArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspPro 141  
Db 2037 CGAGGACACCCAGCTCTTGGAGACACATATGCCCTGGCGGTGATGAAATAGAGACCCG 2096

QY 142 LeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeuGlnLeu 161  
Db 2097 CTGAACCATTACCCCTGTGTACAGGGGCTCCCCAGAGGCTGGGGAGCTGCACCTT 2156

QY 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAspProGlnLeuGly 181  
Db 2157 CGAAGCCCTCACAGAGATTTTGAAGAGGGGTCTTGATCCAGCGGACCCCACTGTCGC 2216

QY 182 TyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201  
Db 2217 TACAGAGACAGATTTTGTGGAAGACATCTCCACAGAACAACACACTGGCTGTACA 2276

QY 202 LeuIleAspThrAsnArgSerArgAlaLysHisProCysSerProMetCysLysGlySer 221  
Db 2277 CTGATAGACACCAACCGCTCTCGGGCTGCGCACCCCTCTTCTCCGATGTGAAGGCTCC 2336

QY 222 ArgCysTrpGlyGluSerSerGlnAspCysGlnSerLeuThrArgThrValCysAlaGly 241  
Db 2337 CGCTGCTGGGAGAGAGTTCTGAGGATTTGCAGACCTGAGCGGACACTGTCTGTCCGGT 2396

QY 242 GlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAlaAla 261  
Db 2397 GCGTGTGCGCGCTGCAGAGGGGCACTGCCACTGACTCTGCGATGAGACAGTGTCTGCC 2456

QY 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281  
Db 2457 GGTGTGACAGGGGCCCCAACCTGTGACTGTGCTGCGCTGCTCCACTCAACACAGTGGC 2516

QY 282 IleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMet 301  
Db 2517 ATCTGTAGCTGCACCTGCCCGCCCTGTGTCACCTACAAACAGACAGCTTTGATGATCATG 2576

QY 302 ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn 321  
Db 2577 CCCAATCCCCGAGGCGCGGTATACATTCGGCCGACAGTGTGACTGTGCTGCTCCACAAAC 2636

QY 322 TyrLeuSerThrAspValGlySer----- 329  
Db 2637 TACCTTTCTACGAGCGGTGATCCTGCACCTCTGCTGCCCCCTGCACAAACAGAGGTG 2696

QY 329 ----- 329  
Db 2697 ACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCACAGAACCCCTGTGCCGAGTGTGC 2756

QY 329 ----- 329  
Db 2757 TATGCTGTGGCATGAGACACTTGCAGAGGTGAGGGCATTTACCAATATATCCAG 2816

QY 329 ----- 329  
Db 2817 GAGTTTGCTGCTGCAGAGAGATCTTTGGAGACCTGTGCATTCTGCGGAGAGCTTTGAT 2876

QY 329 ----- 329  
Db 2877 GGGGACCCAGCTCCAAACATGCCCCGCTCCAGCCAGACAGCTCCAAAGTGTTGAGACT 2936

QY 329 ----- 329  
Db 2937 CTGGAAGATACAGAGTTTACCTATATCTACAGATGCGCGAGACCTCTGACCTC 2996

QY 329 ----- 329  
Db 2997 AGCCTTTCAGAAACCTGCAGTAATCCGGGAGAAATTCTGCACAAATGCGCCTACTCG 3056

QY 329 ----- 329  
Db 3057 CTGACCCCTGAAAGGCTGGGATCATGAGTGGGGGCTGCGCTCACTGAGGGACTGGGC 3116

QY 329 ----- 329  
Db 3117 AGTGACTGGCCCTCATCACCATTAAACACCACCTCTGCTTCGTGCACACGCTGGCCCTG 3176

QY 329 ----- 329  
Db 3177 GACCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCGAGAGGAC 3236

QY 329 ----- 329  
Db 3237 GAGTGTGTGGGCGAGGGGCTGGGCTGCCACAGCTGTGCGGCCGAGGACACTGTGGGGT 3296

QY 329 ----- 329  
Db 3297 CCAGGGCCACCAAGTGTGTAAGTGCAGCCAGTTCTTCGGGGGCGAGAGTGCCTGGAG 3356

QY 329 ----- 329  
Db 3357 GAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGACACTGTGTTGCCG 3416

QY 329 ----- 329  
Db 3417 TGCCACCTGAGTGCACGCCCGAGAAATGGCTCAGTGAAGTCTTTTGGACCGGAGGCTGAC 3476

QY 329 ----- 329  
Db 3477 CAGTGTGTGGCTGTGCCACTATTAAGAACCTCCCTCTTGTGCTGTGGCCGCTGCCACAGC 3536

QY 329 ----- 329  
Db 3537 GGTGTGAACCTGACCTCTCTATACATGCCCATCTGGAAGTTTCCAGATGAGAGAGGCGCA 3596

329 ----- 329  
3597 TGCCAGCCTTGGCCCATCACTGACCCACCTCTGTGTGACCTGATGACAAAGGCTGC 3656  
329 ----- 329  
3657 CCCGCCAGCAGAGAGCCAGCCCTCTGACGTCATGCTCTGTGGGTTGGCATTCCTG 3716  
329 ----- 329  
3717 CTGCTGCTGTCTTGGGGGTGTCTTTGGATCTCTATCAAGCAGCAGCAGAAAGATC 3776  
329 ----- 329  
3777 CGGAAGTACAGATGCGAGACTGCTCAGAGAAAGAGCTGTGGAGCCGCTGACACT 3836  
329 ----- 329  
3837 AGCGAGCGATGCCCAACAGCGCAGATGCGATCTCTAAAGAGACGAGCTGAGAG 3896  
329 ----- 329  
3897 GTGAAGTGTGTGATGTGGGCTTTTGGCAGCTACAAAGGCTATGATCCCTGAT 3956  
329 ----- 329  
3957 GGGAGATGTGAATAATTCAGTGGCCATCAAGTGTGAGGAAACACATCCCCCAA 4016  
329 ----- 329  
4017 GCCAAGAAAGAAATCTTAGACGAAGCATACGTATGCTGTGGCTGCCCTCATATGTC 4076  
329 ----- 329  
4077 TCCCGCTCTGGGATCTGCTGATCAATCACAGGTGACGTGTGACACAGCTTATGCC 4136  
329 ----- 329  
4137 TATGGCTCTTATGACCATGTCCGGGAAACCGCGAGCGCTGCCAGACTG 4196  
329 ----- 329  
4197 CTGAAGTGTATGACGATTTGCCAAGGGATGATGACTGAGAGATGCGGCTGTA 4256  
329 ----- 329  
4257 CACAGGACTTGGCCGCTCGGAACGTGCTGTAAGAGTCCCAACATGTCAAAATTACA 4316  
329 ----- 329  
4317 GACTTCGGGCTGGCTGGCTGTGACATTGACGAGACAGATACATGAGATGGGGC 4376  
329 ----- 329  
4377 AAGTGGCCATCAAGTGGATGGCGCTGAGTCCATTCTCCGGCGGCTTACCCACCAG 4436  
329 ----- 329  
4437 AGTATGTGTGAGTATGTGTGTGACTGTGTGGAGCTGATGACTTTTGGGCGCAACT 4496  
329 ----- 329  
4497 TACGATGGATCCAGCCGCGGAGATCCCTGACTGTGAAAAGGGGAGCGGCTGCC 4556  
329 ----- 329  
4557 CAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGCAAAATGTGTGATGATGAC 4616  
329 ----- 329  
4617 TCTGAATGTGGCCAGATTCGGGAGTGTGTCTGAATTTCCCGCATGGCCAGGGAC 4676  
329 ----- 329

4677 CCCAGCGCTTGTGTATCAAGATGAGACTTGGGCCAGCCAGTCCCTTGGACAGC 4736  
329 ----- 329  
4737 ACCTTACCGCTCAGTCTGAGAGCAGATGACATGGGGACCTGTGTGATGCTGAGAG 4796  
330 ----- -GLYALAGLYMet 334  
4797 TATGTGTACCCAGCAGGGCTTCTGTCTACAGACCTGCCCCGGCGGTGGGCGATG 4856  
335 ValHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThrLeuGly 354  
4857 GTCCACACAGGACCCAGCTCATACAGAGTGGCGGTGGGAGCTGACACTGAGG 4916  
355 LeuGluProSerGluGluGlyAlaProArgSerProLeuAlaProSerGluGlyAlaGly 374  
4917 CTGAGCCCTCTGAAGAGAGGCCCCAGGCTCTCAGTGGCAGCCCTCCGAAGGGGCTGGC 4976  
375 SerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuPro 394  
4977 TCCGATGATTTGATGTGACCTGGAATGGGGGACCAAGGGGCTGCAAGGCTGCC 5036  
395 ThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeuProSer 414  
5037 ACACATGACCCACCCCTCTACACAGGCTACAGTACAGACCCACAGTACCTGCTCT 5096  
415 GluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAsnGln 434  
5097 GAGACTATGCTACGTACGTGGCCCTGACCTGACGACCCCAAGCTGATATGTAAACAG 5156  
435 ProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAlaArgPro 454  
5157 CCAGATGTGGGCCCCAGCCCTTCCGCCGAGAGGGCCCTGCTGCTGCCGAGCT 5216  
455 AlaGlyAlaThrLeuGluAlaGlyAlaGlyThrLeuSerProGlyLysAsnGlyValValLys 474  
5217 GCTGGTCCACTGTGAAGGGCCAAAGACTCTCCAGGAGAAATGGGCTGTCAA 5276  
475 AspValAlaThrLeuGlyAlaValAlaGlyLysAsnProGluTyrLeuThrProGlnGly 494  
5277 GACCTTTTGGCTTGGGGGCTGCCGTGAGAACCCGAGTACTTGACACCCAGGAGGA 5336  
495 AlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 514  
5337 GCTGCCCTCAGGCCACCCCTCTGCTTCCAGCCAGCCCTTCCGACAACTCTATTAC 5396  
515 TrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThrProThr 534  
5397 TGGACCAAGACCCACAGAGCGGGGCTCCACCCAGCACTTCAAAAGGACACCTTAC 5456  
535 AlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSerPro 554  
5457 GCAGAGAACCCAGAGTACCTGGGTCTGAGCTGCA ----- GTCGTACACAGAGGCCA 5510  
555 SerProSerThrGlnProThrGlnHisValAlaAlaLeuGlnGluAlaArgArgLeu 574  
5511 AGTCCCGAG--AAAGCCCTATGTGTCTCAGG-----GAGCAGGGAAGGGCGCTCTG 5561  
575 AsnLeuSerArgAspThr 580  
5562 AGCTATTCCAGAAAGTAGT 5579

RESULT 12  
AR080259 4473 bp DNA Linear PAT 31-AUG-2000  
LOCUS AR080259  
DEFINITION Sequence 1 from patent US 5968748.  
ACCESSION AR080259  
VERSION AR080259.1 GI:10006994  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 4473)  
 AUTHORS Bennett, C. Frank, L. Lipton, A. and Witters, L. M.  
 TITLE Antisense oligonucleotide modulation of human HER-2 expression  
 JOURNAL Patent: US 5968748-A 1 19-OCT-1999;  
 FEATURES Location/Qualifiers  
 source 1. 4473  
 BASE COUNT 902 a 1383 c 1329 g 859 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,25e-80 Length: 4473  
 Score: 2418.00 Matches: 523  
 Percent Similarity: 41.06% Conserved: 5  
 Best Local Similarity: 40.67% Mismatches: 22  
 Query Match: 64.67% Indels: 736  
 DB: Gaps: 4  
 US-09-821-883-2 (1-690) x AR080259 (1-4473)  
 QY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17  
 Db 118 CGCCGCCCTCCAGCCGGGCTCCAGCCGACCATGGGGCCGGAGCCGACGTAGCACCATG 177  
 QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValIleuAlaLysGluLeuAlaArg 37  
 Db 178 GAGCTGGGGCGCTTGCGCGCTG-----GGGCTCTCTCTGCGCCCTCTTGCCCCC 228  
 QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57  
 Db 229 GGAGCGCGCAGCACCACCAAGTGTGCACCGCGCACAGACATGACCTGCGCTCCCTGCACAT 288  
 QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGly 77  
 Db 289 CCGGAGACCCACTGCTGACATGCTCCGCGCACCTCTACACAGGCTGCCAGGTGGTCAGAGGA 348  
 QY 78 AsnLeuGluLeuThrTyrlLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97  
 Db 349 AACCTGGAACTCACCCTACTGCGCCACCAATGCCAGGCTCTCTCTCTGAGGATATCCAG 408  
 QY 98 GluValGlnGlyTyrlValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117  
 Db 409 GAGGTGCAGGGCTAGCTGCTCATCGCTCACAACCAAGTGAAGGAGGTCCCTGCTGACAGAG 468  
 QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrlAlaLeuAlaValLeuAsp 137  
 Db 469 CTGGCGATTGTGCGAGGACCCAGCTCTTTGAGGACACTATGCCCCGCTGGCTGAGAC 528  
 QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157  
 Db 529 AATGGAGACCGGCTGAAACAATACCAACCCCTGTACAGGGGCTCCCGCAGAGAGCCGTGGG 588  
 QY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsn 177  
 Db 589 GAGCTGCACCTTCGAAGCTTCACAGAGACTTGAAGAGGGGCTTGTGATCCAGCGGAAAC 648  
 QY 178 ProGlnLeuCysTyrlGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln 197  
 Db 649 CCCCAGCTCTGCTACAGGACACAGATTTTGTGAGAGACATCTTCCACAGAAACAACCCAG 708  
 QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPromet 217  
 Db 709 CTGGGCTCTACACTGATATACACCAACCCCTCTCGGGCTTGCACCCCTGTCTCGGATG 768  
 QY 218 CysLysGlySerArgCysTrpLysGluSerSerGluAspCysGlnSerLeuThrArgThr 237  
 Db 769 TGTAAAGGCTCCCGCTGCTGGGAGAGAGTCTTGAGGATTTGTACAGAGCTTGACGCGCACT 828  
 QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu 257  
 Db 829 GTCTGGCGGGGTGGTGTGTGCTGCAAGGGGCCACTCCCACTGACTGTGCTGCCATGAG 888  
 QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277

Db 889 CAGTGTGCTGCCGGCTGCAGGGGCCCAAGCACTGACTGCTGCTGCTGCCTCCACTTC 948  
 QY 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrlAsnThrAspThr 297  
 Db 949 AACCAAGAGGAGTGTGAGAGTGCATGCGCCAGCGCTGCTACCTACCAACAACAGACAG 1008  
 QY 298 PheGluSerMetProAsnProGluGlyArgTyrlThrPheGlyAlaSerCysValThrAla 317  
 Db 1009 TTTGAGTCCATCCCAATCCCAAGGCGGCTTACATTCGGCGCCAGCTGTGACTGCC 1068  
 QY 318 CysProTyrlAsnTyrlLeuSerThrAspValGlySer----- 329  
 Db 1069 TGTCCCTACAACTACCTTTCTACGAGACGTGGATCTCCTGACCCCTGTGCCCTGCAC 1128  
 QY 329 ----- 329  
 Db 1129 AACCAAGAGTGCAGACAGAGATGGAACACAGCGGTGTGAGAAGTGCAGAACCCCTGT 1188  
 QY 329 ----- 329  
 Db 1189 GCCCGAGTGTGCTATGCTGTGGGCATGAGACACTTCCGAGAGTGAAGGCACTTACACT 1248  
 QY 329 ----- 329  
 Db 1249 GCCAATATCCAGAGATTGCTGCTGCTGCAGAGAGATCTTTGGAGCCTGSCATTTCTGCCG 1308  
 QY 329 ----- 329  
 Db 1309 GAGAGCTTTGATGGGAGCCAGCCCTCCAAACACTGCCCCCTCCAGCAGCAGCTCCAA 1368  
 QY 329 ----- 329  
 Db 1369 GTGTTGAGACTCTGAGAGAGATCACAGGTTACTTATACATCTCAGCATGCGCGACAGC 1428  
 QY 329 ----- 329  
 Db 1429 CTGCTGACCTCAGGCTTTTCCAGAACTGCAAGTAATCCGGGAGCAATTTCCACAAAT 1488  
 QY 329 ----- 329  
 Db 1489 GCGGCTACTGCTGACCTGCAAGGCTGAGGCTGAGCTGCTGCGGCTGCGCTCACTG 1548  
 QY 329 ----- 329  
 Db 1549 AGGGAACCTGGCAGTGGACTGGCCCTCATCCACATACACCCACTCTGCTTCTGTGCAC 1608  
 QY 329 ----- 329  
 Db 1609 ACGGTGCCCTGGGACCAGCTTTTGGAAACCGGCACCAAGCTCTGCTCCACACTGCCAAC 1668  
 QY 329 ----- 329  
 Db 1669 CGGCAGAGAGCAGTGTGTGGCGAGGCGCTGGCCCTGCACCACTGTGCGCCGAGGG 1728  
 QY 329 ----- 329  
 Db 1729 CACTGCTGGGGTTCAGGGGCCACCCACAGTGTTCACACTGCAGCCAGTTCCTTGGGGGCAG 1788  
 QY 329 ----- 329  
 Db 1789 GAGTGGTGAGAGAAATGCCAGTACTGACAGGGGCTCCCGAGGAGATGATGATGACAGG 1848  
 QY 329 ----- 329  
 Db 1849 CACTGTTTGCCTGGCACCCCTAGTGTACAGCCCAAGATGCTCACTGACTGTTTGA 1908  
 QY 329 ----- 329  
 Db 1909 CCGGAGGCTGACACAGTGTGTGCTGTGCCCACTATAGGACCTTCCCTTGTGCGTGGCC 1968  
 QY 329 ----- 329

Db 1969 CGCTGCCCAAGGCTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2028  
QY 329 ----- 329  
Db 2029 GAGGAGGGGCGATGCCAGCTTGCCCATCACTGCACCCACTCTCTGTGGACCTGGAT 2088  
QY 329 ----- 329  
Db 2089 GACAAAGGCTGCCCCCGGAGAGAGAGCCAGCCCTCTGACGTCCATCATCTGCGGTG 2148  
QY 329 ----- 329  
Db 2149 GTTGGCATTTCTGCTGCTGCTGTCTTGGGGGTGGCTTTTGGGATCCTTCATCAAGCAGG 2208  
QY 329 ----- 329  
Db 2209 CAGAGAAAGATCGGGAAGTACAGATCGGAGACTGCTGCAGAAAGGAGCTGTGGAG 2288  
QY 329 ----- 329  
Db 2269 CCGCTGACACCTAGCGGAGCGATGCCCMACGCGCAGATCGGATCTGAAAGAGACG 2348  
QY 329 ----- 329  
Db 2329 GAGCTGAGGAAGGTGAAGGTGCTTGATCTGCGCTTTTGGCACAGTCTACAAAGGCATC 2388  
QY 329 ----- 329  
Db 2389 TGGATCCCTGATGGGAGAAATGTGAAATTCAGGTGCGCATCAAAAGTGTGAGGAAAC 2448  
QY 329 ----- 329  
Db 2449 ACATCCCCCAAGCCAAAGAAATCTTAGACGAGACATACGTAGTGGTGTGGG 2508  
QY 329 ----- 329  
Db 2509 TCCCATATGTCTCCCGCTTCTGGGCACTGCTGTCAGATCCAGGTGCACTGTGTACA 2568  
QY 329 ----- 329  
Db 2569 CAGCTTATGCCCTATGCTGCTCTTATAGACCATGTCCGGAAAAACCCGGAGCGCTGGG 2628  
QY 329 ----- 329  
Db 2629 TCCAGAGACTGCTGAAGCTGTGTATGACAGATTGCCAAGGGGATGACCTACCTGGAGAT 2688  
QY 329 ----- 329  
Db 2689 GTGGGCTGTACACAGGACTTGGCGGCTCGGAACGTGCTGTCAAGAGTCCCAACCAT 2748  
QY 329 ----- 329  
Db 2749 GTCAAAATTACAGACTTCGGGCTGGCTGGCTGTGACATTGACGAGAGAGATACCAT 2808  
QY 329 ----- 329  
Db 2809 GCAGATGGGGGCAAGTGGCCATCAAGTGAAGTGGCTGAGTCCATTCTCCCGCGGG 2868  
QY 329 ----- 329  
Db 2869 TTCACCCACAGAGTATGTGTGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTT 2928  
QY 329 ----- 329  
Db 2929 GGGGCCAAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACTGTGAAAAAGGG 2988  
QY 329 ----- 329  
Db 2989 GAGCGGCTGCCAGCCGCCCATCTGCACCATTGATGTACATGATCATGTCAAAATGT 3048  
QY 329 ----- 329  
Db 3049 TGGATGATTGACTCTGAAATGCGGCCAAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCGC 3108

QY 329 ----- 329  
Db 3109 ATGGCAGGGACCCCCAGCGCTTTGTGTCATCCAGATGAGACTTGGGCCAGCCAGT 3168  
QY 329 ----- 329  
Db 3169 CCGTTGGACAGACCTTCTACCGCTCACTGCTGAGAGCATGACATGGGGGACCTGTG 3228  
QY 330 ----- 330  
Db 3229 GATGCTAGAGATATCTGTAACCCAGAGGCTTCTTGTCTCCAGACCTGCCCGGGC 3288  
QY 331 AAGllylmetValHisHisArgHisArgSerSerThrArgSerGlyGlyAsp 350  
Db 3289 GCTGGGGCATGTGTCACACAGGACACCGCAGCTCATCTACAGAGATGGGGGGGAGC 3348  
QY 351 LeuThrLeuGlyLeuGlyProSerGlyGlyGlyAlaProArgSerProLeuAlaProSer 370  
Db 3349 CTGACACTAGGGCTGGAGCCCTTGAAAGAGAGGCCCCAGGTCTCCATGCGCACCTCC 3408  
QY 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyAlaLeuGlyLeu 390  
Db 3409 GAAAGGCTGCTGCTCCGATGATTTGATGCTGACCTGGGAATGGGGCCAGCCAGGCTG 3468  
QY 391 GluSerLeuProThrHisAspProSerProLeuGlyAlaArgTyrSerGlyAspProThrVal 410  
Db 3469 CAAAGCTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAAGACCCACAGTA 3528  
QY 411 ProLeuProSerGlyThrArgGlyTyrValAlaProLeuThrGlySerProGlyProLeu 430  
Db 3529 CCGTGCCTCTGAGACTGATGCTACCTTGGCCCTTACCTGACCTGACCCCAAGCTGAA 3588  
QY 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyGlyProLeuPro 450  
Db 3589 TATGTGAACACAGCAGATGTTCGGCCCGCCCTTGGCCGAGAGGGCCCTCTGCT 3648  
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGlyAlaAlaLysThrLeuSerProGlyLysAsn 470  
Db 3649 GCTGCCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3708  
QY 471 GlyValAlaLysAspValPheAlaPheGlyGlyValAlaGlyLysProGlyLeuThr 490  
Db 3709 GGGCTGCTCAAGAGCTTTTGGCTTGGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3768  
QY 491 ProGlnGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510  
Db 3769 CCCAGGGAGAGCTGCCCTCAGCCCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3828  
QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGlyAlaAlaProProSerThrPheLys 530  
Db 3829 AACCTTATTAATCTGGGACACAGGACCCAGAGGGGGGCTCCACACAGCCTTCAAA 3888  
QY 531 GlyThrProThrAlaGlyAsnProGlyTyrLeuGlyLysAspValProAlaAlaPro 550  
Db 3889 GGCACACTACGCGCAGAGAACCCAGAGTACCTGGGTCTGACGTGCA -----GTGTGA 3942  
QY 551 AlaArgSerProSerPro 556  
Db 3943 ACCAGAGGCCAAGTCCG 3960  
RESULT 13  
AR167390 4473 bp DNA linear PAT 17-DEC-2001  
LOCUS AR167390  
DEFINITION Sequence 26 from patent US 6287569.  
ACCESSION AR167390  
VERSION AR167390.1 GI:17903168  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 4473)  
AUTHORS Kipps,T.J. and Wu,Y.

TITLE Vaccines with enhanced intracellular processing  
JOURNAL Patent: US 6287569-A 26 11-SEP-2001;  
FEATURES Location/Qualifiers  
SOURCE 1. 4473  
/organism="unknown"  
BASE COUNT 902 a 1383 c 1329 g 859 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.25e-80 Length: 4473  
Score: 2418.00 Matches: 523  
Percent Similarity: 41.06% Conservative: 5  
Best Local Similarity: 40.67% Mismatches: 22  
Query Match: 64.67% Indels: 736  
DB: 6 Gaps: 4

US-09-821-883-2 (1-690) x AN167390 (1-4473)

QY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17  
DB 118 CGCGCCCTCCCGACCGGGGTCCAGCCGAGCCATGGGGCCGAGCCGAGTGAACCATG 177  
QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37  
DB 178 GAGCTGGCGGCTTGTGCGCGTGG-----GGGCTCTCTCTGCGCCCTCTGCCCCC 228  
QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57  
DB 229 GGAGCCGGACACCAACCAAGTGTGACCGCGACAGACATGAAGCTCGGCTCCCTCCAGT 288  
QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGluValValGlnGly 77  
DB 289 CCCGAGACCAACCTGAGATGCTCGCCACACCTTACAGAGGCTGCGAGTGTGACAGGA 348  
QY 78 AsnLeuGluLeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97  
DB 349 AACCTGGACATCACTTACTGCTCCACCAATGCCACTGCTCTCTCTCAGAGATATCCAG 408  
QY 98 GluValGlnGlyTrpValLeuLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117  
DB 409 GAGGTGACGGGCTACGTGCTCTACCTCACAACCAAGTGAAGGACAGGTCCCACTGACAG 468  
QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAsp 137  
DB 469 CTGGCGATTTGGCGAGGACCAACCACTTTGAGGACAATAGCCCTGGCGGTGTAGAC 528  
QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157  
DB 529 AATGAGACCGCGTGAACATACCAACCCCTGTACAGGGGCTCCCAAGAGGCTTGGCG 588  
QY 158 GluLeuGlnLeuArgSerLeuThrGluLeuLysGlyValLeuIleGlnArgAsn 177  
DB 589 GAGCTGACAGCTTCGAAAGCTTCACAGAGATCTTGAAAGGGGGCTTTGATCCAGCGAAG 648  
QY 178 ProGlnLeuCysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGln 197  
DB 649 CCCAGCTCTGCTACAGAGACAGATTTTGGAGAGACATCTCCACAGAACAAACACAG 708  
QY 198 LeuAlaLeuThrIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217  
DB 709 CTGGCTTCACACTGATGAGACCAACCGCTCTCGGGCTGACACCCCTGTTCTCCGATG 768  
QY 218 CysLysGlySerArgCysTrpGlyLysSerGlnAspCysGlnSerLeuThrArgThr 237  
DB 769 TGTAAAGGCTCCCGCTGCTGGGAGAGAGTTCTGAGATTGTACAGACCTGACGGCACT 828  
QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGln 257  
DB 829 GTCTGTCCGCGTGTGCGCGTGCAGAGGGGCCACTGCCACTGCTGCGCATGAG 888  
QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277  
DB 889 CAGTGTGCTGCGGCTGACAGGGGCCCAAGCACTCTGACTGCTGCGCTGCCACTTCACATTC 948

QY 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThr 297  
DB 949 AACCAACATGGGATCTGTGAGCTGACCTGACCTGCCAGCCCTGCTACATACACAGACAGCAG 1008  
QY 298 PheGluSerMetProAsnProGlnGlyArgTrpThrPheGlyAlaSerCysValThrAla 317  
DB 1009 TTTGAGTCCATGCCCAATCCGAGGGCGGTATACATTGCGGCGCAGCTGTGTGACTGCC 1068  
QY 318 CysProTrpAsnTrpLeuSerThrAspValGlySer----- 329  
DB 1069 TGTCCCTCAACTACTCTTCTTAACGAGCGTGGGATCTGACCCCTGCTG-CGCCCTGAC 1128  
QY 329 ----- 329  
DB 1129 AACCAAGAGTGCACAGAGAGATGGAACACAGCCGCTGAGAACTGCACAGCAAGCCCTGT 1188  
QY 329 ----- 329  
DB 1189 GCCCGAGTGTGCTATGCTGTGGGATGAGACACTTGGCAGAGGTGAGGGCAGTTACAGT 1248  
QY 329 ----- 329  
DB 1249 GCCAATATCCAGAGTTTGTGCTGCAAGAAGATCTTTGGAGCGCTGGCATTTCTGCCG 1308  
QY 329 ----- 329  
DB 1309 GAGACCTTGTATGGGGGACCCAGCCTCCAAACACTGCCCGGCTCCAGACAGACAGCTCCA 1368  
QY 329 ----- 329  
DB 1369 GTTTTGAAGACTGTGGAAGAGATACAGAGTTACTATATCATCTAGCATGGCCGAGACAG 1428  
QY 329 ----- 329  
DB 1429 CTGCTGACCTCAGCGCTTCTCCAGAACTGCAGAAATCCGGGAGCAATTCGACAAAT 1488  
QY 329 ----- 329  
DB 1489 GGGCGCTACTGCTGACCTGCAAGGCTGGGACATCAGCTGGCTGGGCTGGCTGCTCAGTG 1548  
QY 329 ----- 329  
DB 1549 AGGGAAGTGGGAGTGTGAGTGGCCCTCATCCACATTAACCCACCTGTGCTTCTGAC 1608  
QY 329 ----- 329  
DB 1609 ACGGTGCTGGGACCACTTTTGGAAACCCGACCAAGCTGTGCTCCACACTGCCAAG 1668  
QY 329 ----- 329  
DB 1669 CGGCCAGAGAGAGTGTGTGGGCGAGGGCTGGGCTGCCACCACTGTGCCCGGAGAG 1728  
QY 329 ----- 329  
DB 1729 CACTGCTGGGTCCAGGGCCCAACCAAGTGTGCAACTGACGACCACTTCTGGGGCGAG 1788  
QY 329 ----- 329  
DB 1789 GAGTGTGAGAGAAATGCCGAGTACTGCAAGGGCTCCCGAGGAGATATGTGAATGCCAG 1848  
QY 329 ----- 329  
DB 1849 CACTTTTCCCGTGCACCTGAGTGTGACAGCCCAAGATGGCTCAGTCACTGTTTGA 1908  
QY 329 ----- 329  
DB 1909 CCGAGGCTGACAACTGTGTGGCTGTGCCACTATTAAGAACCTCCCTTGTGCTGCGCC 1968  
QY 329 ----- 329  
DB 1969 CGTGTCCCAAGCGGTGAACCTGACCTTCTCTACATGATGCCACTGTGAAGTTTCCAGAT 2028

QY 329 ----- 329  
Db 2029 GAGAGGGCGCATGCGACGCTTGCCCATCACTGACACCCACTCTGTGGACCTGGAT 2088  
QY 329 ----- 329  
Db 2089 GACAAAGGCTGCCCCCGGACAGACAGACCACTCTGACGCTCATTCTCTGCGGTG 2148  
QY 329 ----- 329  
Db 2149 GTTGGCATTCGTCGTGCTGTGGGGGTGCTTTGGGATCCTCATCAACGACGG 2208  
QY 329 ----- 329  
Db 2209 CAGCAGAAGATCCGGAAGTACACAGTCGAGACTGCTGAGAAAGCGACTGTGGAG 2268  
QY 329 ----- 329  
Db 2269 CCGCTGACACCTAGCGAGGAGATGCCAACAGCGCATGCGGATCTGAAAGAGACG 2328  
QY 329 ----- 329  
Db 2329 GAGCTGAGAAAGTGAAGTGTGATCTGGCGCTTTTGGCACAGTCTCAAGGCGATC 2388  
QY 329 ----- 329  
Db 2389 TGGATCCCTGATGGGAGAAATGTGAAATTCACGTGCGCATCAAAAGTTGAGGAAAC 2448  
QY 329 ----- 329  
Db 2449 ACATCCCCCAAGCAACAAAGAAATCTTACAGCAACATACGTGATGGCTGGTGGGC 2508  
QY 329 ----- 329  
Db 2509 TCCCATATGTCTCCGCTTCTGGGCATCTGCTGACATCCAGCGTGCAGCTGTGACA 2568  
QY 329 ----- 329  
Db 2569 CAGCTATGCCCTATGGCTGCTCTTACACCATGTCGCGAAMACCAGCGACGCTGGGC 2628  
QY 329 ----- 329  
Db 2629 TCCAGAGACCTGCTGAAGTGTGATGACATGTCACAAAGGGGATGACTACTGGAGAT 2688  
QY 329 ----- 329  
Db 2689 GTGGGCTCTGACAGAGGACTTGGCGCTCGAAGCTGCTGTCAGAGTCCCAACCAT 2748  
QY 329 ----- 329  
Db 2749 GTCAAAATTACAGACTTCGGGGCTGGCTGCTGTGACATTGACGAGAGAGACTACAT 2808  
QY 329 ----- 329  
Db 2809 GCAGATGGGGGCAAGTGGCCATCAAGTGAAGCGCTGAGTCAATTCCCGCGCGG 2868  
QY 329 ----- 329  
Db 2869 TTCACCACCAAGATGATGTGTGATTATGTGTGACTGTGTGAGAGTGAATCTTTT 2928  
QY 329 ----- 329  
Db 2929 GGGGCCAAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACGTGAGAAAGGG 2988  
QY 329 ----- 329  
Db 2989 GAGCGGCTGCCCAAGCCCATCTGCACCATTTGATGTATACATGATGATGTCAAATGT 3048  
QY 329 ----- 329  
Db 3049 TGGATGATGACTGTGAATGTGGCCAAAGATTCGGGAGTTGGTGTGAATTCCTCCGC 3108  
QY 329 ----- 329

Db 3109 ATGCCAGAGACCCCGACGCTTTGTGTCATCCAGATGAGACTTGGGCCAGCCAGT 3168  
QY 329 ----- 329  
Db 3169 CCCTTGGACAGCACCTTCTACCGCTCAGCTGTGGAGACATGACATGGGAGACTGTG 3228  
QY 330 ----- 330  
Db 3229 GATGCTGAGAGATATCTGTATCCAGAGGCGCTTCTGTCTCCAGACCTGCCCGGCG 3288  
QY 331 ----- 331  
Db 331 AAGAGTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 3368  
QY 3389 GCTGGGGGAGTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 3448  
QY 351 ----- 351  
Db 3349 CTGACACTAGGAGCTGAGACCTCTGAAGAGAGAGGCGCCAGCTCCACCTGACCTCC 3408  
QY 371 ----- 371  
Db 3409 GAAAGGCTGCTGCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 3468  
QY 391 ----- 391  
Db 3469 CAAAGCCTCCGACATGAGACCCGACGCTCTACAGCGGATGAGAGACCCACAGTA 3528  
QY 411 ----- 411  
Db 3529 CCGCTGCTGCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 3588  
QY 431 ----- 431  
Db 431 TATGTGAACAGCAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 3648  
QY 451 ----- 451  
Db 3649 GCTGCCGAGCTGCTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 3708  
QY 471 ----- 471  
Db 3709 GGGGCTGCTCAAGACGTTTTCCTTGGGGTCCCTGGAGAACCCGAGTACTTACA 3768  
QY 491 ----- 491  
Db 3769 GGGGCTGCTCAAGACGTTTTCCTTGGGGTCCCTGGAGAACCCGAGTACTTACA 3828  
QY 511 ----- 511  
Db 3829 AACCTTATTAAGTGGAGACAGACAGACAGAGCGGGGGCTCCACACAGACCTTCAA 3888  
QY 531 ----- 531  
Db 3889 GGGGCTGCTCAAGACGTTTTCCTTGGGGTCCCTGGAGAACCCGAGTACTTACA 3942  
QY 551 ----- 551  
Db 3943 ACCAGAGGCCAAGTCCG 3960  
RESULT 14  
HSEB2R  
LOCUS HSEB2R 4473 bp mRNA linear PRI 30-MAR-1995  
DEFINITION Human c-erb-B-2 mRNA.  
ACCESSION X03363.1 GI:31197  
VERSION X03363.1  
KEYWORDS cell surface glycoprotein; cellular oncogene; erb-2 cellular;  
glycoprotein; growth factor receptor; kinase; neu cellular  
oncogene; transmembrane protein; tyrosine kinase.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Mammalia: Eutheria: Chordata: Craniata: Vertebrata: Euteleostomi;  
1 (bases 1 to 4473)



**AUTHORS** Yamamoto, T., Ikawa, S., Akiyama, T., Semba, K., Nomura, N., Miyajima, N., Saito, T. and Toyoshima, K.  
**TITLE** Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor  
**JOURNAL** Nature 319 (6050), 230-234 (1986)  
**MEDLINE** 86118663  
**PUBMED** 3003577  
**REFERENCE** 2 (bases 1 to 4473)  
**AUTHORS** Papewalis, J., Nikitin, A.Yu. and Rajewsky, M.F.  
**TITLE** G to A polymorphism at amino acid codon 655 of the human erbB-2/HER2 gene  
**JOURNAL** Nucleic Acids Res. 19 (19), 5452 (1991)  
**MEDLINE** 92020265  
**PUBMED** 1681519  
**COMMENT** The c-erb-B-2 protein shows similarity to the epidermal growth factor receptor.  
**FEATURES** Location/Qualifiers  
**source** 1..4473  
 /organism="Homo sapiens"  
 /strain="cell line MKN-7"  
 /db\_xref="taxon:9606"  
 175..3942  
 /note="c-erb-B-2 precursor"  
 /codon\_start=1  
 /protein\_id="CA27060.1"  
 /db\_xref="GI:31198"  
 /db\_xref="SWISS-PROT:P04626"  
 /translation="MELALACRMGLLALLLPFGAASRQVCTGDMKRLRPSPTHLDMRLYOGCQVQVGNLELYLPNNAISLFQDIDVGGYLIANNQVQVQLRLRYRGTLQFDNVALAVLNDGDLNNTPTVPGASPGRLRELQLSLRELKSGVLDLRNQLCYQDITLMKDIHFKNQMLATLIDTNSRCGSPKSGRCMGSESECCSLTRTVACAGCARGKGLPTDCCHECAGCTGPKHSDDLACHENHSGICELHCPALTYNTDTEPMNPEGRYFEGASCYATPYNTLDSCTLYACPLAHNEVTAEGCTORCECSKPRACVCIQIGMHLREVRAYTSANTDFRAGCKITGSLAFPESTFDGASNTALQPEQVFEILEETIGLYLSAMPDLSLDFEQLDYLKRLIHNAYSLTDLGLISWLCRLRELKSGSLIHNHTLCEVHYVWDLFRNPHQALHTANRDECVGSLACHQLARGCHGMPPTQVNCISQFLRGQCEVECVLQGLPREYVNAKICLCHPECQPNQSVTCGEPADQVCANHKDPFCVARGPSGVKPLSYNPLMKPFDEGACQPCINCHSCDLDKGCAPAOBASPLTISAVVGLLVVGVGIIIRKROOKIRKYMRLLOTELEVERPLTSGAMPNQAOKILKETELRKVKVIGSGAFGVYKGTINLPGENKIPVAIKVLRNTSKANKETLDEYIVAGVSPVSRLLGICLTSTVQLVTPQMGKGLDHNRENRGLSQDLLNMCQIAKMSYLEDRVLVHRDLAANVLYKSPNHVKTDFGLARLIDIDETEHADGKVPYIKWALLESILRRRFHOSDWSYGVTVELMTFARGPYDGIIPAREIPDLLEKGERLPOPICTIDYVIMIKCMIDSECRPFRELVSERSMARQRFVIVIONEDIGPASPLDSFYSLEDDMDGLVAEEYLPQOGFPCBPARGAGMVAHRRHSSTRSGCDLTLGPSEEARPSRLASSEGASVDFGDLGMAKKGDSLTPHPSLPQRTSEDTVPLESETGTYAPLTCSFQPEYVNPQVAPQPPSPREGPLPAARPAAGATLERKTLSPKNGVYKVFARFGAVENPEYLTTPGGGAAPOHPPPAFSPAFDNLVYWDQDPPERGAPSTFEGKPTAEINPEYLGIDVPY"  
 175..237  
 /note="putative"  
 238..3939  
 /product="put. c-erb-B-2 protein (aa 1-1234)"  
 376..384  
 /note="pot. glycosylation site"  
 544..558  
 /note="pot. glycosylation site"  
 733..741  
 /note="pot. glycosylation site"  
 949..957  
 /note="pot. glycosylation site"  
 1762..1770  
 /note="pot. glycosylation site"  
 1885..1893  
 /note="pot. glycosylation site"  
 2039..2067  
 /note="pot. glycosylation site"  
 2353..3132  
 /note="aa 727-986, seq. homologous to EGF receptor kinase domain"  
 2446..2454  
 /note="pot. glycosylation site"  
 4435..4460  
 misc\_feature  
 misc\_feature  
 misc\_feature

polyA\_site  
 BASE COUNT 902 a 1383 c 1329 g 859 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,25e-80 Length: 4473  
 Score: 2418.00 Matches: 523  
 Percent Similarity: 41.06% Conservative: 5  
 Best Local Similarity: 40.67% Mismatches: 22  
 Query Match: 64.67% Indels: 736  
 DB: Gaps: 4  
 US-09-821-883-2 (1-690) x HSEBR2R (1-4473)  
 QY 2 ArgAlaAlaProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17  
 DB 118 CGCGCCCTCCAGCGCGGTCCAGCGGAGCCATGGGGCCGAGCGGAGTGCACCATG 177  
 QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37  
 DB 178 GAGCTGGCGGCGCTTGCGCGCTGG-----GGCTCTCTCCGCGCCCTTGCCCCCC 228  
 QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57  
 DB 229 GGAGCCGCGAGCACCACCAAGTGTGCACCGGCACAGCATGAGCTCGCTCCCAT 288  
 QY 58 ProGluTrpHisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGly 77  
 DB 289 CCCGAGACCACACCTGACATGCTCGCGCACCTCAGACGGGCTGCAGGTGGTCAGGGA 348  
 QY 78 AsnLeuGlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97  
 DB 349 AACCTGGAACTCCTACCTGCTGCCACCAATGCACTGCTCTCTCAGGATATCCAG 408  
 QY 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117  
 DB 409 GAGGTGCAGGGCTAGCTGCTCATGCTCACAACCAAGAGGAGGAGTCCACACTGCAGAGG 468  
 QY 118 LeuArgGlnLeuValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137  
 DB 469 CTGCGGATGTCGCGAGGACACCGACTTGTGAGGCAACATAGCCCTGCTGCTACAG 528  
 QY 138 AsnGlnAspProLeuAsnAsnThrThrProValThrGlnAlaSerProGlnGlyLeuArg 157  
 DB 529 AATGAGAGCCCGCTGAACAATACACACCCCTGTCAAGAGGAGGCTCCCAAGAGGCTGCGG 588  
 QY 158 GluLeuGlnLeuValArgSerLeuThrGlnLeuLysGlnGlyValLeuIleGlnArgAsn 177  
 DB 589 GAGCTGCAGGCTTCGAAAGCTCACAAGATCTTGAAAGAGGGGCTTGTGATCCAGCGGAAC 648  
 QY 178 ProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGln 197  
 DB 649 CCCACGCTCTGCTACAGACACGATTTGTGGAGGAGCACTTCCACAAAGAACACAG 708  
 QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217  
 DB 709 CTGGCTCTACACTGATATGACACCAACCGCTCTCGGGCTGCGACCCCTGTTCTCGCATG 768  
 QY 218 CysLysGlySerArgCysTrpLysLysSerGlnAspCysGlnSerLeuThrArgThr 237  
 DB 769 TGTAGAGGCTCCCGCTGCTGGGGAAGAAGTCTGAGATTTGACAGAGCTGACGGCACT 828  
 QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGln 257  
 DB 829 GTCCTGCGCGGCTGCGCGCGTGCAGAGGGCCACATGCCATGCTGCTGCTCATAG 888  
 QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277  
 DB 889 CAGTGTGCTGCGGCTGCAGCGGCGCCCAAGACACTGTACTGCTGCTGCTCCATTC 948  
 QY 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297

Db	944	AACCCACAGTGCATCTGTGAGCTGCACCTGCCCGCCCTGGTCACCTACAAACACAGACAGC	1008
QY	238	PheGluSerMetProAsnProGluGlyAlaGlyTyrThrPheGlyAlaSerCysValThrAla	317
Db	1009	TTTGAGTCTCAATGCCCAATCCCCGAGGGCCGGTATATACATTCGGGCCAGAGTGTGACTGCC	1068
QY	318	CysProTyrAsnTyrLeuSerThrAspValGlySer	329
Db	1069	TGTCTCTACAACTACTCTTTCTTACGGAGCGTGGATCTCTGACCCCTGTCGCCCCCTGCAC	1128
QY	329	-----	329
Db	1129	AACCAAGAGTGCACAGACAGAGATGGAACACAGCGGTGTGAGAAGTGCAGAACGCTGT	1188
QY	329	-----	329
Db	1189	GCCCGAGTGTCTATWGTGTCTGGCCATGAGACACTTGCAGAGAGTGAGGGCAGTTACCACT	1248
QY	329	-----	329
Db	1249	GCCAAATATCCAGAGATTGCTGGCTGCAGAAAGATCTTTGGAGAGCCTGGCATTTCTGCCG	1308
QY	329	-----	329
Db	1309	GAGAGCTTTGATGGGGAGACCCAGCCTCCAAACTGCCCCGCTCCAGACAGAGAGCTCCA	1368
QY	329	-----	329
Db	1369	GTTGTTGAGACTCTGGAAGAGATCACAGGTTACTATATCATCTCAGCATGGCCGGACAGC	1428
QY	329	-----	329
Db	1429	CTGCCTGACCTCAGACCGTCTTCCAGAAACCTGCAAGTAAATCCGGGGAGCAATTCGCACAA	1488
QY	329	-----	329
Db	1489	GGCGCTACTGCTGACCTCGAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCACTG	1548
QY	329	-----	329
Db	1549	AGGGAAGTGGGCACTGGACTGGCCCTTCATCCACCAATAAACCCACCTTGTCTGTGCAC	1608
QY	329	-----	329
Db	1609	ACGGTCCCTGGAGACAGCTCTTTCGGAACCCGACCAAGCTCTGTCCACACTGCCAAC	1668
QY	329	-----	329
Db	1669	CGGCCAGAGACGAGTGTGTGGGCGAGGGCTGGCTGCCACAGCTGTGGCCCGAGGG	1728
QY	329	-----	329
Db	1729	CACGTGTGGGGTCCAGGGGCCACCAAGTGTGTAACTGAGCAGATTCTCTGGGGGCCAG	1788
QY	329	-----	329
Db	1789	GAGTGCCTGAGAGAAATGCCGAGTACTGCAGGGGCTCCCGAGGAGATATGTGAATGCCAG	1848
QY	329	-----	329
Db	1849	CACGTGTTGCCGCTGCCACCTGTAGGTCAAGCCCAAGATGGCTCAGTGAACCTGTTTGA	1908
QY	329	-----	329
Db	1909	CCGGAGGCTGACCAAGTGTGTGGCTGTGCCCACTAATAAGAACCTCCCTTCTGGCTGGCC	1968
QY	329	-----	329
Db	1969	CGCTGCCCAAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTCCAGAT	2028
QY	329	-----	329

Db	2029	GAGGAGGGCCGATCCAGCGCTTGCCCATCACTGACCCACCTCTGTGTGGACTGGAT	2088
QY	329	-----	329
Db	2089	GACAAAGGCTGCCCCCGCAGACAGAGACGACGCCCTGTAGCTCATCTCTGCGGTG	2148
QY	329	-----	329
Db	2149	GTTGGCATTTGCTGTGTCGTGCTTTGGGGGTGTCCTTTGGATCCTCATCAAGCAGCG	2208
QY	329	-----	329
Db	2209	CAGCAAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGAGAAACGAGCTGTGAG	2268
QY	329	-----	329
Db	2269	CCGCTGACCTTACGGAGCGATGCCCAACAGCGCAGATCGGATCCTGMAAGACG	2328
QY	329	-----	329
Db	2329	GAGCTAGGAAGTGGAAGTGCTTGATCTGCCGCTTTTGGCAGCTACAGAGGCATC	2388
QY	329	-----	329
Db	2389	TGGATCCCTGATGGGGAGATGTGAAATTCAGTGGCATCAAGTGTGAGGAAAC	2448
QY	329	-----	329
Db	2449	ACATCCCCCAAGCCAAAGAAATCTTAGAGGAAGCATACGTGTGCTGTGCGC	2508
QY	329	-----	329
Db	2509	TCCCCATATGTCTCCGCCCTTCTGGGCATCTGCTGACATCCAGGTCAGCTGTGACA	2568
QY	329	-----	329
Db	2569	CAGCTATGCCCTTGGCTGCCCTTAGACATGTCCGGAAACCGGAGCCTGGGC	2628
QY	329	-----	329
Db	2629	TCCAGACCTGTGACTGTGTATGACAGATTGCCAAGGAGTAGCTACCTGAGAGT	2688
QY	329	-----	329
Db	2689	GTCGGGCTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGTCCAAACAT	2748
QY	329	-----	329
Db	2749	GTCAAATTAAGACTTCGGGGTGGGCTCGGCTGCTGGACATTGACAGAGACAGATACAT	2808
QY	329	-----	329
Db	2809	GCAGTGGGGCAGGTGCCCATCAAGTGAATGCGCTGAGTCCATTCTCGCGCGCG	2868
QY	329	-----	329
Db	2869	TTTCAACCAAGAGTATGTGTGAGCTTATGTGTGACTGTGTGGAGCTGATCACTTTT	2928
QY	329	-----	329
Db	2929	GGGGCAACCTTACGATGGGATCCCAAGCCGGGAGATCCCTGACCTGCTGMAAAGGG	2988
QY	329	-----	329
Db	2989	GAGGGGTGCCCGACCCCCCATCTGACACATTGATGTCTACATGATCATGTCAATGT	3048
QY	329	-----	329
Db	3049	TGGATGATGACTCTGAATGTGCGCCAGATTCGGGAGTGTGTGAATTTCTCCGC	3108
QY	329	-----	329
Db	3109	ATGGCCAGGACCCCCAGCCCTTTGTGTGATTCACAGATGAGACTTGGGCTCCAGCCAGT	3168

QY	329	-----	329
Db	3169	CCCTTGACAGACACTTCTACCGCTCAGTGCAGATGACATGAGGGGACCTGGTG	3228
QY	330	-----Gly	330
Db	3229	GATGCTGAGGATATCTGGTACCCAGAGAGGGGCTTCTGTGTCAGACCTGGCCCCGGGC	3288
QY	331	AlaGlyGlyMetValHisHisIAsrGlySerSerSerThrArgSerGlyGlyLysasp	350
Db	3289	GCTGGGGGACATGCTCCACACAGCAGCCAGCTCATCTACAGAGATGGGGTGGGAGAC	3348
QY	351	LeuThrLeuGlyLeuGluProSerSerGluGluGluAlaProArgSerProLeuAlaProSer	370
Db	3349	CTGACACTAGAGGCTGAGACCCCTGTGAAGAAGGGGCCCCAGGTCTTCCACTGGCACCTTCC	3408
QY	371	GluGlyAlaGlySerAspValPheAspGlyLysAspLeuGlyMetGlyAlaAlaLysGlyLeu	390
Db	3409	GAAGGGGCTGGGCTCCGATGATTTATGTGTGACCTGGGAATGGGGGACGCCAAGGGGGTG	3468
QY	391	GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal	410
Db	3469	CAAACCTCTCCACACATGACACCCAGCCCTCTACAGCGGTACAGTACAGAGACCCACAGTA	3528
QY	411	ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu	430
Db	3529	CCCCGCCCCCTGACACTGATGTGCTACGTTGCCCCCTGACCTGACAGCCCCAGCCTGA	3588
QY	431	TyrValAsnGlnProAspValArgProGluInProProSerProArgGluGlyProLeuPro	450
Db	3589	TATGTGAACACACAGATGTTGGCCCGACGCCCTTGCAGCCCGAGAGGGCCCTGTGCT	3648
QY	451	AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn	470
Db	3649	GCTGCCGACCTGCGGTGGCCACTGTGGAAGGCCCAAGACTCTCTCCACAGGAAGAT	3708
QY	471	GlyValValIlyAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr	490
Db	3709	GGGGTGGTCAAAAGAGCTTTTGTCTTGGGGGTGGCGTGGAGAACCCGAGTACTTGACA	3768
QY	491	ProGlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp	510
Db	3769	CCCCAGGGGAGGCGTCCCTTCAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3828
QY	511	AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys	530
Db	3829	AACCTCTATTACTGGAGCAGGACCCACACAGCGGGGGGCTCCACCCAGCACCTTCAAA	3888
QY	531	GlyThrProThrAlaGluAsnProGluTyrLeuGlyLysAspValPheAlaAlaAlaPro	550
Db	3889	GGGACACTACGCGCAGAGAACCCACAGTACCTGGGTCTGGAGCTGTCCA-----GTGTGA	3942
QY	551	AlaArgSerProSerPro 556	
Db	3943	ACCAGAGGGCCAAAGTCCG 3960	
RESULT 15			
AX060704			
LOCUS	AX060704	3768 bp	DNA
DEFINITION	Sequence 2 from Patent WO0100244.		
ACCESSION	AX060704		
VERSION	AX060704.1	GI:12406101	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 3768)		
TITLE	Erickson, S. and Schwall, R.		
	Methods of treatment using anti-erbB antibody-maytansinoid		
	conjugates		
JOURNAL	Patent: WO 0100244-A 2 04-JAN-2001;		

FEATURES		Genentech, Inc. (US)	
SOURCE	LOCATION/Qualifiers		
	1. .3768		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	758 a 1170 c 1121 g 719 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3.59e-80	Length:	3768
Score:	2410.50	Matches:	510
Percent Similarity:	41.16%	Conservative:	0
Best Local Similarity:	41.16%	Mismatches:	2
Query Match:	64.47%	Indels:	727
DB:	6	Gaps:	1
US-09-821-883-2 (1-630) x AX060704 (1-3768)			
QY	35	LeuAlaArgGlyAlaAlaSerThrGlnValCysThrGlyThrAspMetCysLeuArgLeu	54
DB	46	TTGGCCCCCGGAGCGCGGAGACCCAAAGTGTGGACACGGGACAGACATGAAGCTGGCGCTC	105
QY	55	ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnVal	74
DB	106	CGTGCAATCCCGAAGACCACCTGGACATGCTCCGCCACCTCTAACCGAGGCTCCAGGTG	165
QY	75	ValGlnGlyAsnLeuGlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGln	94
DB	166	GTGCAGGGGAACCTGGACACTACCTACTGCCACCAATGCAAGCCTGTCTCTCCAG	225
QY	95	AspIleGlnGluValGlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValPro	114
DB	226	GATATCCAGAGAGGTCCAGGCTAGCTGTCTCATCCCTCAACCAACAACTGAGCGAGTCCCA	285
QY	115	LeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAla	134
DB	286	CTGCAGAGCTGGCGATTGTGTGGAGCACCCAGCTCTTTGAGCAACATATGCCCTGGCC	345
QY	135	ValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGly	154
DB	346	GTGTCTAGACAATGGAGACCCCTGAAACAAATCCACCCCTGTCAACAGGGGCTCCCAAGA	405
QY	155	GlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIleLeuGlyValGlyAlaIle	174
DB	406	GGCTCTCGGAGACTGCACACTTCGAAGCCTCAACAAGATCTTGAAAGAGGGGCTTGTATC	465
QY	175	GlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeuThrPlyAspIlePheHisLys	194
DB	466	CAGCGGAACCCCAAGCTGTGTACAGAGACACGATTTGTGTGAAGGACATCTTCCACAG	525
QY	195	AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCys	214
DB	526	AACAACACAGCTGGCTCTACACTGATTAACACCAACCGCTTCGGGCTCCCAACCTCTGT	585
QY	215	SerProMetCysLysGlySerArgCysTyrPylIleuSerSerGlyAspCysGlnSerLeu	234
DB	586	TCTCCGATGTGAAGGCTCCCGCTGCGGGGAAGAGATTGTAGAGATTGTCAAGCTG	645
QY	235	ThrArgThrValCysAlaGlyIleCysAlaArgCysLysGlyProLeuProThrAspCys	254
DB	646	ACGGCACCTGTCTGTGGCGGTGGCTGCCCGGTGCAAAAGGGGACATGCCACATGACTGC	705
QY	255	CysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys	274
DB	706	TGCCATGGACAGTGTGTGGCGGTGGACGGGCCCCAAGACATCTGTACTCCCTGGCCCTGC	765
QY	275	LeuHisPheAsnHisSerGlyIleCysGluLeuHisLysCysProAlaLeuValIleThrTyrAsn	294
DB	766	CTCCACCTTCAACACAGTGGCATGTGTGAGCTGTGACACGTGCCACGCTCGTCACTAAC	825
QY	295	ThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCys	314
DB	826	ACAGACAGCTTTGAGTCCATCCCAATCCGAGGGGCGGTATACATTCCGCGCCAGCTGT	885

0y	315	VaThPAlAcysPrCyYrAsnTYLLeuSerThAspAlGlySer-----	329
Db	886	GTGACTGCTGTCCCTACCACTACTCTTTCAGSAGCTGGATCCGACCCCTGCTGC	945
0y	329	-----	329
Db	946	CCCCGCACAAACAAGAGGTGACGACGAGGATGGAACACAGCGGTGTGAGAAATGACAC	1005
0y	329	-----	329
Db	1006	AAGCCCTGTCCCGAGTGTGCTATAGTCTGTGGCATGAGACACTTGGAGAGGTGAGGCA	1065
0y	329	-----	329
Db	1066	GTTTACCACTGCCAATATCCAGAGAGTTTGTCTGCTGCAGAAAGATCTTTTGGAGCCTGGCA	1122
0y	329	-----	329
Db	1126	TTTCTGCCGAGAGACTTTGATGGGAGCCAGCCTCCAAACACTGTCCCCGCTCCAGCCAGAG	1185
0y	329	-----	329
Db	1186	CAGCTCCAAGTGTTTGAGACTCTGGAAGAGATCAGAGGTTACTTATACATTCCAGACATGG	1245
0y	329	-----	329
Db	1246	CCGACACGCTGCTGACCTCAGCGTCTTCCAGAACCTGCAAGTAAATCCGGGGAGCAATT	1305
0y	329	-----	329
Db	1306	CTGCACAATGGCGCCTACTGCTGACCTTGCAAGGCGTGGGCAATCAGCTGTGGGGCTG	1365
0y	329	-----	329
Db	1366	CGCTCACTGAGAGGAAGTGGGCAGTGGAGCTGGCCCTCATCCACCATACACACCACCTTGC	1425
0y	329	-----	329
Db	1426	TTCTGCACACGAGTGCCCTGGGACACACTTTCGGAAACCCGACCAACACTCTGCTGCAC	1485
0y	329	-----	329
Db	1486	ACTGCCAACCGGCCACAGAGACAGAGTGTGTGGGAGAGGGCCTGGCTGCCACCAAGCTGTGC	1545
0y	329	-----	329
Db	1546	GCCCGAGGGCACACTGTGGGGTCCAGGGCCACCCACAAGTGTCAACTGACAGCCAGTTCCCT	1605
0y	329	-----	329
Db	1606	CGGGGCCAAGAGTGCCTGGAGAGAAATGCCAGTACTGCAGGGGCTCCCCAGGAGATATGTG	1665
0y	329	-----	329
Db	1666	AATGCACAGCACACTGTTCGCCGTGCACACTGAGTGTACAGCCCAAGAAATGGCTCAGTGAC	1725
0y	329	-----	329
Db	1726	TGTTTTGACCGAGAGCTGCACAGTGTGTGGCTGTGCCCACTATAAGAGACCCTCCCTTC	1785
0y	329	-----	329
Db	1786	TGCGGCGCCCGCTGCCCCAGCGGTGTAAACCTGACCTCTCTACATGTGCCCATCTTGGAA	1845
0y	329	-----	329
Db	1846	TTTCAGATGAGAGGGCGCATGCCACCTTGGCCCATCAACTGCACCCACTCTGTGTG	1905
0y	329	-----	329
Db	1906	GACCTGATGACAAAGGCGTGGCCCCCGGACAGACAGAGCCAGCGCTCTGACGTCCATCGTC	1965

QY	329	-----	329	-----	329
Db	1966	TCTGGGTTGGCATTTCTGCTGTGTGTCTTGGGGGTGCTTTGGATCCTCATC	2025		
QY	329	-----	329	-----	329
Db	2026	AAGGACGGCAGCAGAAATCCGGAAGTACAGATGCCGAGACTGTGCAGGAAACGGAG	2085		
QY	329	-----	329	-----	329
Db	2086	CTGTGGAGCCGCTGACACCTAGCGGAGGATGCCACAGGCGCAGATCGGATCCTG	2145		
QY	329	-----	329	-----	329
Db	2146	AAAGAGCGGAGCTGAGGAAGTGAAGTGCTTGGATCTGGGCTTTTGGCAGAGTGTAC	2205		
QY	329	-----	329	-----	329
Db	2206	AAGGCATCTGGATCCTGTATGGGAGAAATGTGMAAATTCAGTGGCCATCAAGTGTG	2265		
QY	329	-----	329	-----	329
Db	2266	AGGMAAACACATCCCCCAAGCCAAAGAAATCTTAGAGAAAGCATAGTGTGCT	2325		
QY	329	-----	329	-----	329
Db	2326	GGTGTGGGCTCCCATATGTCTCCGCTTGTGGCATCTGCCGTGACATCCACGGTGCAG	2385		
QY	329	-----	329	-----	329
Db	2386	CTGTGACACAGCTTATGCCCTATGGCTGCCTTTAGACCATGTCCGGAAAAACGCGGA	2445		
QY	329	-----	329	-----	329
Db	2446	CGCCTGGCTCCAGACCTGTGAATGTGTATGATGACATTGCCAAGGATGAGTAC	2505		
QY	329	-----	329	-----	329
Db	2506	CTGGAGATGTGGGCTGTGACACAGGACTTGGCCGCTCGGACGTGCTGGTCAAGACT	2565		
QY	329	-----	329	-----	329
Db	2566	CCCAACCATGTCAAAATTTACAGACTTCGGGCTGGCTCGGCTGTGGACATTGACGAGACA	2625		
QY	329	-----	329	-----	329
Db	2626	GAGTACCATGAGATGGGGGCAAGTGCCCATCAAGTGAATGCGCTGAGTCCATTCTC	2685		
QY	329	-----	329	-----	329
Db	2686	CGCCGGCGGTTCAACCCACAGAGTGAATGTGAGTTATGTTGTACTGTGTGGAGCTG	2745		
QY	329	-----	329	-----	329
Db	2746	ATGACTTTTGGGCAAACTTACAGATGGGATCCAGCCCGGAGATCCCTGACTGTCTG	2805		
QY	329	-----	329	-----	329
Db	2806	GAAGAGGGGAGCGGCTGCCCAAGCCCCCATCTGCACATTGATGTCTACATGATCATG	2865		
QY	329	-----	329	-----	329
Db	2866	GTCAAAATTTGATGATTGACTGTGAATGTGCGCCAAAGATTCGGGAGTTGTTGTGAA	2925		
QY	329	-----	329	-----	329
Db	2926	TTCTCCGCATGGCCAGGAGACCCCAAGCGCTTGTGTGTCATCAGAAATGAGACTTGGC	2985		
QY	329	-----	329	-----	329
Db	2986	CCAGCCAGTCCCTTGGAGACAGACTTCTACCGCTCACTGCTGGAGAGAGATGACATGGG	3045		
QY	329	-----	329	-----	329

Db 3046 GACCTGTGATGCTGAGAGTATCTGTACCCAGAGGGCTTCTGTCTCAGACCT 3105  
QY 330 -----GlyAlaGlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGly 347  
Db 3106 GCGCCGGGGCTGGGGATGCTGCACACAGGACCGCAGCTCATCTACAGGAGTGGC 3165  
QY 348 GlyGlyAspLeuThrLeuGlyLeuGlyProSerGlyGlyAlaProArgSerProLeu 367  
Db 3166 GGTGGGGACCTGACACTGAGGCTGGAGCCCTCTGAAGAGAGGCCCCAGGCTTCCACTG 3225  
QY 368 AlaProSerGlyGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAla 387  
Db 3226 GCACCTCCGAAGGGGCTGGCTCCGATGTATTGTATGTGACCTGGGAATGGGGCAGCC 3285  
QY 388 LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGlyAsp 407  
Db 3286 AAGGGCTGCAAGGCTCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGAGAC 3345  
QY 408 ProThrValProLeuProSerGlyThrAspGlyTyrValAlaProLeuThrCysSerPro 427  
Db 3346 CCCACAGTACCCCTGCTCCCTGAGAGCTATGCTACGTTGCCCTGACCTGCAGCCCC 3405  
QY 428 GlnProGlyTyrValAsnGlnProAspValArgProGlnProProSerProArgGlyGly 447  
Db 3406 CAGCCTGAATATGTGAAACACAGCATGTGCGCCCCAGCCCCCTTCGCCCGAGAGGGC 3465  
QY 448 ProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlyAlaGlyThrLeuSerPro 467  
Db 3466 CCTGCTGCTGCTGCCGACCTGCTGTCCTGCACTGTGAAAGGGCCAAAGACTCTCTCCCA 3525  
QY 468 GlyLysAsnGlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGlu 487  
Db 3526 GGGAAAGATGGGGTGTCTCAAGACGTTTGTGCTTTGGGGGTGCCGTGAGAACCCCGAG 3585  
QY 488 TyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerPro 507  
Db 3586 TACTTGACACCCAGGAGGAGAGCTGCCCTCAGCCCACTCTGCTTCAGCCCA 3645  
QY 508 AlaPheAspAsnLeuTyrTyrTyrAspGlnAspProProGlnArgGlyAlaProProSer 527  
Db 3646 GCGTTCGACAACTCTATCTAGGACAGAGACCCAGAGCGGGGGCTCCACCCAGC 3705  
QY 528 ThrPheLysGlyThrProThrAlaGluAsnProGlyTyrLeuGlyLeuAspValPro 546  
Db 3706 ACCTTCAAAGGACACCTACGGCAGAGAACCCAGAGTACTGGTCTGAGAGTGCCA 3762

Search completed: May 8, 2003, 14:16:48  
Job time : 4411 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:40:24 ; Search time 25.3846 Seconds  
(without alignments)  
2178.074 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739  
Sequence: 1 MRAAPLLARAASLSGLFLF.....EPVOGAPPPAAHHHHHH 690

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCR\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3739	100.0	690	10 US-09-821-883-2	Sequence 2, Appl1
2	3473.5	92.9	697	10 US-09-821-883-4	Sequence 4, Appl1
3	2972	79.5	555	10 US-09-821-883-1	Sequence 1, Appl1
4	2957.5	79.1	564	10 US-09-821-883-3	Sequence 3, Appl1
5	2573.5	68.8	919	9 US-09-834-356-6	Sequence 6, Appl1
6	2410.5	64.5	1255	10 US-09-811-123-9	Sequence 2, Appl1
7	2410.5	64.5	1255	10 US-09-811-115-3	Sequence 1, Appl1
8	2410.5	64.5	1255	9 US-09-854-356-1	Sequence 1, Appl1
9	2405.5	64.3	1255	9 US-09-930-125-2	Sequence 2, Appl1
10	2405.5	64.3	1255	9 US-09-441-411-6	Sequence 5, Appl1
11	2405.5	58.5	479	10 US-09-821-883-5	Sequence 11, Appl1
12	1961	52.4	1256	9 US-09-870-759-14	Sequence 14, Appl1
13	1950.5	52.2	1256	9 US-09-854-356-2	Sequence 2, Appl1
14	1948	52.1	1256	10 US-09-821-161-1	Sequence 1, Appl1
15	1612	43.1	653	9 US-09-854-356-3	Sequence 3, Appl1
16	1612	43.1	712	9 US-09-854-356-7	Sequence 7, Appl1
17	1612	43.1	289	10 US-09-821-883-23	Sequence 23, Appl1
18	1587	42.4			
19					

20	1367.5	36.6	654	9 US-09-854-356-8	Sequence 8, Appl1
21	1183	31.6	266	9 US-09-854-356-4	Sequence 4, Appl1
22	1183	31.6	583	9 US-09-930-125-9	Sequence 9, Appl1
23	1183	31.6	587	9 US-09-930-125-8	Sequence 8, Appl1
24	1183	31.6	589	9 US-09-930-125-10	Sequence 10, Appl1
25	1183	31.6	600	9 US-09-930-125-11	Sequence 11, Appl1
26	1182	31.6	217	10 US-09-821-883-25	Sequence 25, Appl1
27	1182	31.6	397	10 US-09-821-883-27	Sequence 27, Appl1
28	1182	31.6	1179	10 US-09-821-883-29	Sequence 29, Appl1
29	893	23.9	191	9 US-09-441-411-9	Sequence 9, Appl1
30	721	19.3	657	9 US-10-172-620-18	Sequence 18, Appl1
31	721	19.3	1210	10 US-09-725-433-2	Sequence 2, Appl1
32	721	19.3	1308	10 US-09-940-101-2	Sequence 2, Appl1
33	717	19.2	293	9 US-10-102-806-583	Sequence 583, App
34	713.5	19.1	615	10 US-09-940-101-4	Sequence 4, Appl1
35	704.5	18.8	478	10 US-09-867-521-2	Sequence 2, Appl1
36	697	18.6	1342	9 US-10-172-620-16	Sequence 16, Appl1
37	684	18.3	610	10 US-09-783-708-1	Sequence 1, Appl1
38	674	18.0	144	10 US-09-923-246-114	Sequence 114, App
39	673	18.0	127	9 US-09-792-793A-15	Sequence 15, Appl1
40	673	18.0	127	10 US-09-821-883-18	Sequence 18, Appl1
41	673	18.0	127	10 US-09-800-016-1	Sequence 11, Appl1
42	648.5	17.3	135	10 US-09-925-301-1232	Sequence 1232, Ap
43	441	11.8	127	10 US-09-821-883-20	Sequence 20, Appl1
44	373.5	10.0	219	10 US-09-847-185-2	Sequence 2, Appl1
45	255	6.8	120	9 US-10-172-620-17	Sequence 17, Appl1

## ALIGNMENTS

RESULT 1  
US-09-821-883-2  
Sequence 2, Application US/09821883  
Patient No. US20020061310A1  
GENERAL INFORMATION:  
APPLICANT: Laus, Reiner  
APPLICANT: Vidovic, Damir  
APPLICANT: Graddis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821, 883  
CURRENT FILING DATE: 2001-03-30  
PRIORITY APPLICATION NUMBER: US 60/193,504  
PRIORITY FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 690  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500-HCM-CSF construct  
US-09-821-883-2  
Query Match 100.0%; Score 3739; DB 10; Length 690;  
Best Local Similarity 100.0%; Pred. No. 4e-212;  
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MRAAPLLARAASLSGLFLFELFWLDRSVLAKELARGAASVOYCTGDMKRLPASPEPT 60  
DB 1 MRAAPLLARAASLSGLFLFELFWLDRSVLAKELARGAASVOYCTGDMKRLPASPEPT 60  
OY 61 HMDMRLHYOGCGVOVGNLELTYLPTNASLSFLDIOEYGVYLAHNOVROPYLORLRI 120  
DB 61 HMDMRLHYOGCGVOVGNLELTYLPTNASLSFLDIOEYGVYLAHNOVROPYLORLRI 120  
OY 121 VAGTOLFEDNVALAVLDNGDPLNNTPTVGTASPGGLRELOJLSRLREIKGVLIORNPOL 180  
DB 121 VAGTOLFEDNVALAVLDNGDPLNNTPTVGTASPGGLRELOJLSRLREIKGVLIORNPOL 180  
OY 181 CYODTILWKDIFHKNNOLATFLIDITNRSRACHPCSPMKSGRCNCESSDCSLTRTYCA 240

```

Db 181 CYDDTILMKDIFKHNKNOALATLIDTNRSRACHCSPKCSRCWGESSEDCOSLTRVCA 240
OY 241 GGCARCKGRLPTDCHCQCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
Db 241 GGCARCKGRLPTDCHCQCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
OY 301 MPNDEGRYTFGASCVTACPNYVLTSDVSGAGGVNHHRRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNDEGRYTFGASCVTACPNYVLTSDVSGAGGVNHHRRSSSTRSGGDLTLGLEPSEE 360
OY 361 EARRSPPLABEGASDVFDGDLGMAAGKQSLPTHPDPSLQRYSEDPYVPLSEDTGYV 420
Db 361 EARRSPPLABEGASDVFDGDLGMAAGKQSLPTHPDPSLQRYSEDPYVPLSEDTGYV 420
OY 421 APLTCSQPEYVNOVDVRRPSPREGPLPAPRAGATLERAATLSPGKGVKDVAFG 480
Db 421 APLTCSQPEYVNOVDVRRPSPREGPLPAPRAGATLERAATLSPGKGVKDVAFG 480
OY 481 GAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLVYWDODPERGAPSPSTFKGTPAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLVYWDODPERGAPSPSTFKGTPAENPEY 540
OY 541 LGIDVPAARASPSSTOQWENHNAIOEARLLNSRDTAAENETVEYISEMFDLOEP 600
Db 541 LGIDVPAARASPSSTOQWENHNAIOEARLLNSRDTAAENETVEYISEMFDLOEP 600
OY 601 TCLQTRLELYKQGLRSLTKLGPLTMMASHYKONCPPTETSCATOIITFESEKLENKD 660
Db 601 TCLQTRLELYKQGLRSLTKLGPLTMMASHYKONCPPTETSCATOIITFESEKLENKD 660
OY 661 FLVLPEDCMEPVQEGAPPPAAAHNNHH 690
Db 661 FLVLPEDCMEPVQEGAPPPAAAHNNHH 690

```

```

RESULT 2
US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damlir
; APPLICANT: Gradalis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-tgm-CSF construct
US-09-821-883-4

```

```

Query Match 92.9%; Score 3473.5; DB 10; Length 697;
Best Local Similarity 91.7%; Pred. No. 1.6e-196;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;

```

```

OY 1 MRAAPLLAARASLSIGFLFFLFWLDRSVLAKELARGASTOVCTGDMKLRLPASPER 60
Db 1 MRAAPLLAARASLSIGFLFFLFWLDRSVLAKELARGASTOVCTGDMKLRLPASPER 60
OY 61 HLDMLHLYOGCGVOGNLELTLYLPTNASLSFODIOEVGYVLIHNOVROYVPLQRLRI 120
Db 61 HLDMLHLYOGCGVOGNLELTLYLPTNASLSFODIOEVGYVLIHNOVROYVPLQRLRI 120

```

```

OY 121 VRGTQLEEDNYALAVLNDGDLPLNNTTPVTGASPGGLREQLRSLTEILKGVYLQORPOL 180
Db 121 VRGTQLEEDNYALAVLNDGDLPLNNTTPVTGASPGGLREQLRSLTEILKGVYLQORPOL 180
OY 181 CYDDTILMKDIFKHNKNOALATLIDTNRSRACHCSPKCSRCWGESSEDCOSLTRVCA 240
Db 181 CYDDTILMKDIFKHNKNOALATLIDTNRSRACHCSPKCSRCWGESSEDCOSLTRVCA 240
OY 241 GGCARCKGRLPTDCHCQCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
Db 241 GGCARCKGRLPTDCHCQCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
OY 301 MPNDEGRYTFGASCVTACPNYVLTSDVSGAGGVNHHRRSSSTRSGGDL 351
Db 301 MPNDEGRYTFGASCVTACPNYVLTSDVSGAGGVNHHRRSSSTRSGGDL 351
OY 352 TLGLEPSEEEARPSPLASGAGSDVDGDLGMAAGKQSLPTHPDPSLQRYSEDPYV 411
Db 352 TLGLEPSEEEARPSPLASGAGSDVDGDLGMAAGKQSLPTHPDPSLQRYSEDPYV 411
OY 412 LPSETDGYVAPLTCSPQPEYVNOVDVRRPSPREGPLPAPRAGATLERAATLSPGKNG 471
Db 412 LPSETDGYVAPLTCSPQPEYVNOVDVRRPSPREGPLPAPRAGATLERAATLSPGKNG 471
OY 472 VKRDVEAFGAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLVYWDODPERGAPSPSTFKG 531
Db 472 VKRDVEAFGAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLVYWDODPERGAPSPSTFKG 531
OY 532 IPTAENPEYLGIDVPAARASPSSTOQWENHNAIOEARLLNSRDTAAENETVEYI 591
Db 532 IPTAENPEYLGIDVPAARASPSSTOQWENHNAIOEARLLNSRDTAAENETVEYI 591
OY 592 SEMFDLOEPTCLQTRLELYKQGLRSLTKLGPLTMMASHYKONCPPTETSCATOIITFE 651
Db 592 SEMFDLOEPTCLQTRLELYKQGLRSLTKLGPLTMMASHYKONCPPTETSCATOIITFE 651
OY 652 ESFEKLENKDPLVLPEDCMEPVQEGAPPPAAAHNNHH 690
Db 652 ESFEKLENKDPLVLPEDCMEPVQEGAPPPAAAHNNHH 690

```

```

RESULT 3
US-09-821-883-1
; Sequence 1, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damlir
; APPLICANT: Gradalis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct
US-09-821-883-1

```

```

Query Match 79.5%; Score 2972; DB 10; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.6e-167;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MRAAPLLAARASLSIGFLFFLFWLDRSVLAKELARGASTOVCTGDMKLRLPASPER 60
Db 1 MRAAPLLAARASLSIGFLFFLFWLDRSVLAKELARGASTOVCTGDMKLRLPASPER 60

```



QY 61 HLDMLRHLHYGCGVYVGNLELTLYPTNASTLSFLDDIOEVQGYVLIANQVROVPLQRLRI 120  
 DB 61 HLDMLRHLHYGCGVYVGNLELTLYPTNASTLSFLDDIOEVQGYVLIANQVROVPLQRLRI 120  
 QY 121 VRGTQLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRLNQL 180  
 DB 121 VRGTQLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRLNQL 180  
 QY 181 CYQDTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESEDDQSILTRVCA 240  
 DB 181 CYQDTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESEDDQSILTRVCA 240  
 QY 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300  
 DB 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300  
 QY 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVNHHRSSSTRSGGDLTLGLEPSEE 360  
 DB 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVNHHRSSSTRSGGDLTLGLEPSEE 360  
 QY 361 EAPRSPLASRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSTDTGY 420  
 DB 361 EAPRSPLASRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSTDTGY 420  
 QY 421 APLTCSPOPEYVNPQVPRQPPSPREGPLPAARPGATLERAKTILSPKNGVVKDVFAG 480  
 DB 421 APLTCSPOPEYVNPQVPRQPPSPREGPLPAARPGATLERAKTILSPKNGVVKDVFAG 480  
 QY 481 GAVENPEYLTPOGGAAPQHPPPAFSPAFDNLYYWDDPBERGAPSTFKGTPAENPEY 540  
 DB 481 GAVENPEYLTPOGGAAPQHPPPAFSPAFDNLYYWDDPBERGAPSTFKGTPAENPEY 540  
 QY 541 LGLDVPAAP 549  
 DB 541 LGLDVPAAP 549

RESULT 4  
 US-09-821-883-3  
 : Sequence 3, Application US/09821883  
 : Patent No. US20020061310A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Laus, Retner  
 : APPLICANT: Vidovic, Damir  
 : APPLICANT: Gradovic, Thomas  
 : TITLE OF INVENTION: Compositions and Methods for Dendritic  
 : FILE REFERENCE: 7636-0022.30  
 : CURRENT APPLICATION NUMBER: US/09/821,883  
 : CURRENT FILING DATE: 2001-03-30  
 : PRIOR APPLICATION NUMBER: US 60/193,504  
 : PRIOR FILING DATE: 2000-03-30  
 : NUMBER OF SEQ ID NOS: 30  
 : SOFTWARE: FastSeq for Windows Version 4.0  
 : SEQ ID NO 3  
 : LENGTH: 564  
 : TYPE: PRT  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: HERS500\* construct  
 US-09-821-883-3

Query Match 79.1%; Score 2957.5; DB 10; Length 564;  
 Best Local Similarity 98.4%; Pred. No. 2.6e-16;  
 Matches 549; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

DB 61 HLDMLRHLHYGCGVYVGNLELTLYPTNASTLSFLDDIOEVQGYVLIANQVROVPLQRLRI 120  
 QY 121 VRGTQLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRLNQL 180  
 DB 121 VRGTQLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQRLSLTEILKGVLIQRLNQL 180  
 QY 181 CYQDTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESEDDQSILTRVCA 240  
 DB 181 CYQDTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESEDDQSILTRVCA 240  
 QY 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300  
 DB 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300  
 QY 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVNHHRSSSTRSGGDLTLGLEPSEE 360  
 DB 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVNHHRSSSTRSGGDLTLGLEPSEE 360  
 QY 361 EAPRSPLASRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSTDTGY 420  
 DB 361 EAPRSPLASRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVPLPSTDTGY 420  
 QY 421 APLTCSPOPEYVNPQVPRQPPSPREGPLPAARPGATLERAKTILSPKNGVVKDVFAG 480  
 DB 421 APLTCSPOPEYVNPQVPRQPPSPREGPLPAARPGATLERAKTILSPKNGVVKDVFAG 480  
 QY 472 VVKDVFAGVAVENPEYLTPOGGAAPQHPPPAFSPAFDNLYYWDDPBERGAPSTFKG 531  
 DB 472 VVKDVFAGVAVENPEYLTPOGGAAPQHPPPAFSPAFDNLYYWDDPBERGAPSTFKG 531  
 QY 532 TPTAENPEYLGIDVPAAP 549  
 DB 532 TPTAENPEYLGIDVPAAP 549  
 QY 541 TPTAENPEYLGIDVPAAP 558  
 DB 541 TPTAENPEYLGIDVPAAP 558

RESULT 5  
 US-09-854-356-6  
 : Sequence 6, Application US/09854356  
 : Patent No. US2002017567A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Cheever, Martin A.  
 : APPLICANT: Gheysen, Dirk  
 : APPLICANT: Corixa Corporation  
 : APPLICANT: SmithKline Beecham Biologicals S. A.  
 : TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 : FILE REFERENCE: 014058-009810PC  
 : CURRENT APPLICATION NUMBER: US/09/854,356  
 : CURRENT FILING DATE: 2001-05-09  
 : PRIOR APPLICATION NUMBER: US 09/493,480  
 : PRIOR FILING DATE: 2000-01-28  
 : PRIOR APPLICATION NUMBER: US 60/117,976  
 : PRIOR FILING DATE: 1999-01-29  
 : NUMBER OF SEQ ID NOS: 26  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 6  
 : LENGTH: 919  
 : TYPE: PRT  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: Description of Artificial Sequence: fusion protein  
 : OTHER INFORMATION: of ECD and PD of human HBR-2/neu  
 US-09-854-356-6

Query Match 68.8%; Score 2573.5; DB 9; Length 919;  
 Best Local Similarity 56.4%; Pred. No. 1.7e-13;  
 Matches 509; Conservative 0; Mismatches 3; Indels 391; Gaps 1;

QY 1 MRAAPLLARASLSTGLFLFFWLDKSVLAKELARGASTOVCTGDMKRLRPASPET 60  
 DB 1 MRAAPLLARASLSTGLFLFFWLDKSVLAKELARGASTOVCTGDMKRLRPASPET 60  
 QY 61 HLDMLRHLHYGCGVYVGNLELTLYPTNASTLSFLDDIOEVQGYVLIANQVROVPLQRLRI 120  
 DB 61 HLDMLRHLHYGCGVYVGNLELTLYPTNASTLSFLDDIOEVQGYVLIANQVROVPLQRLRI 120

QY 35 LARGAASOVCTGDMKRLRPASPETHDMLRHLHYGCGVYVGNLELTLYPTNASTLSSTQ 94  
 DB 16 LPPGASOVCTGDMKRLRPASPETHDMLRHLHYGCGVYVGNLELTLYPTNASTLSSTQ 75  
 QY 95 DIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPG 154  
 DB 95 DIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPG 154

```

Db 76 DIOEVGVYLIHNOVQVPLQRLRYRGTOLEEDNYALAVLNDGDLNNTPTVTGASPG 135
QY 155 GLRELQRLSTLEILKGVLLQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNSRACHPC 214
Db 136 GLRELQRLSTLEILKGVLLQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNSRACHPC 195
QY 215 SPMCKSRGWESSEDCQSLRTVACAGGACRCKGRLPTDCCHQCAGCTGPRHSDCLAC 274
Db 196 SPMCKSRGWESSEDCQSLRTVACAGGACRCKGRLPTDCCHQCAGCTGPRHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDTFESMPNEGRTTGASCVTACPYNYLSTDVGSCTLYVC 329
Db 256 LHFHNSGICELHCPALVTYNTDTFESMPNEGRTTGASCVTACPYNYLSTDVGSCTLYVC 315
QY 330 ----- 329
Db 316 PLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREAVRANTSANIOEFAGCKKIEGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLQPELOVFELEITGYLYISAMPDLPLDSVFNQVYIGRI 435
QY 330 ----- 329
Db 436 LHNQAVSLTLOGIGISWGLRLSELGSLALIHNTNLFCVHTVPMOQLFRNPQALLH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHQLCARGHMGPRPTQVCNSQFLRGQCEVECHVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHRECOPONGSVTCFGRPADQVACAHYKDRPCVACRPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEEGACQPCPINCCHSCVDLDDKCPARQASPLTSQNEDELAPASPLDSTYRSLLED 675
QY 330 ----- 329
Db 676 DMGDLVDAEBEYLVRQGGFCRDPARQAGMVHNRHSSSTRSGGGLTLGLPSESEEAR 735
QY 364 RSLPLASEGASVDFDGLMGAAKGLQSLPTHDPSLOYSEDPTVYRLSEPTDGYAPL 423
Db 736 RSLPLASEGASVDFDGLMGAAKGLQSLPTHDPSLOYSEDPTVYRLSEPTDGYAPL 795
QY 424 TCSPOPEYVNPQVRPQPSRREGPLPAARPAATLERAKTSLSPGKNVYKDVFAFGAV 483
Db 796 TCSPOPEYVNPQVRPQPSRREGPLPAARPAATLERAKTSLSPGKNVYKDVFAFGAV 855
QY 484 ENBEYLTPOGGAAPQHPAPAFSPARDNLVYMDODPERGAPSTFKGTPTAENPEYLG 543
Db 856 ENBEYLTPOGGAAPQHPAPAFSPARDNLVYMDODPERGAPSTFKGTPTAENPEYLG 915
QY 544 DVP 546
Db 916 DVP 918

```

RESULT 6  
US-09-769-508-2

Sequence 2, Application US/09769508  
Patent No. US20020155527A1  
GENERAL INFORMATION:  
APPLICANT: STUART, SUSAN G.  
APPLICANT: MONAHAN, JOHN J.  
APPLICANT: LANGTON, BEATRICE CLAUDIA  
APPLICANT: HANCOCK, MIRIAM E.C.  
APPLICANT: CHAO, LORINE A.  
APPLICANT: BLUFORD, PETER  
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75  
FILE REFERENCE: BBIO-111-C1  
CURRENT APPLICATION NUMBER: US/09/769,508

```

; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-508-2

Query Match      64.5%; Score 2410.5; DB 9; Length 1255;
Best Local Similarity 41.2%; Pred. No. 9, 1e-134;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LARGASTOYCTGTDKMLRLPASPEHLDMLRHLVGGQVYVGNLELTYLPTNASLSFLQ 94
Db 16 LPPGAASTOYCTGTDKMLRLPASPEHLDMLRHLVGGQVYVGNLELTYLPTNASLSFLQ 75
QY 95 DIOEVGVYLIHNOVQVPLQRLRYRGTOLEEDNYALAVLNDGDLNNTPTVTGASPG 154
Db 76 DIOEVGVYLIHNOVQVPLQRLRYRGTOLEEDNYALAVLNDGDLNNTPTVTGASPG 135
QY 155 GLRELQRLSTLEILKGVLLQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNSRACHPC 214
Db 136 GLRELQRLSTLEILKGVLLQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNSRACHPC 195
QY 215 SPMCKSRGWESSEDCQSLRTVACAGGACRCKGRLPTDCCHQCAGCTGPRHSDCLAC 274
Db 196 SPMCKSRGWESSEDCQSLRTVACAGGACRCKGRLPTDCCHQCAGCTGPRHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDTFESMPNEGRTTGASCVTACPYNYLSTDVGSCTLYVC 329
Db 256 LHFHNSGICELHCPALVTYNTDTFESMPNEGRTTGASCVTACPYNYLSTDVGSCTLYVC 315
QY 330 ----- 329
Db 316 PLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREAVRANTSANIOEFAGCKKIEGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLQPELOVFELEITGYLYISAMPDLPLDSVFNQVYIGRI 435
QY 330 ----- 329
Db 436 LHNQAVSLTLOGIGISWGLRLSELGSLALIHNTNLFCVHTVPMOQLFRNPQALLH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHQLCARRALLSGPTQVCNSQFLRGQCEVECHVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHRECOPONGSVTCFGRPADQVACAHYKDRPCVACRPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEEGACQPCPINCCHSCVDLDDKCPARQASPLTSIVSAVGLLVVVGVEGILT 675
QY 330 ----- 329
Db 676 KRRQOKIRKTYMRLLQETLVEPLTPSGAMPNOAQRILKETELRKVKVLSGAGCTVY 735
QY 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLENTSPRANKELIDEAVVAGVSPYVSRLLGICLSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMYPGCLLDHVENRGRIGLSQDLNMCQIANGMSYLEBVRVYHRDLAARNLVKS 855
QY 330 ----- 329
Db 856 PNHVKTIDFLANLIDIDETEHADGKVPKIMMALESILRRFTHQSDVWSYGVYVWEL 915

```

```

QY 330 ----- 329
Db 916 MTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYIMVCMWIDSECRPRELVE 975
QY 330 ----- 329
Db 976 FSRMARDPQRFVIONEDLGPASPLDSTFYRSLLLEDDMDGLVDAEEXLVPOGFFCPDP 1035
QY 330 --GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 387
Db 1036 APGAGWHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 1095
QY 388 KGLSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPOPEYVNOQDVAPRPPSPREG 447
Db 1096 KGLSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPOPEYVNOQDVAPRPPSPREG 1155
QY 448 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEYLTPOGGAAPRPHPPAFSP 507
Db 1156 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEYLTPOGGAAPRPHPPAFSP 1215
QY 508 AFDNLXYWDODPPERGAPSTFKGPTTAENPEYLGIDVP 546
Db 1216 AFDNLXYWDODPPERGAPSTFKGPTTAENPEYLGIDVP 1254

```

## RESULT 7

US-09-811-123-9

```

: Sequence 9, Application US/09811123
: Patent No. US2002001587A1
: GENERAL INFORMATION:
: APPLICANT: Shaon Erickson
: APPLICANT: Ralph Schwall
: APPLICANT: Mark Sliwowski
: TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EBB
: TITLE OF INVENTION: ANTIBODY-MAYTANSTINOID CONUGATES
: FILE REFERENCE: GENENT.073A2
: CURRENT APPLICATION NUMBER: US/09/811.123
: PRIOR FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/238,327
: PRIOR FILING DATE: 2000-10-05
: PRIOR APPLICATION NUMBER: 09/602,530
: PRIOR FILING DATE: 2000-06-23
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 1255
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-811-123-9

```

```

Query Match 64.5% Score 2410.5; DB 10; Length 1255;
Best Local Similarity 41.2%; Pred. No. 9.1e-134;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

```

```

QY 35 LARGAASOVCTGDMKRLPASPETHLMDLRHLVYCGOVVQGNLELYLPTNASTSLFLQ 94
Db 16 LPPGAASOVCTGDMKRLPASPETHLMDLRHLVYCGOVVQGNLELYLPTNASTSLFLQ 75
QY 95 DIOEVQGVLLAHNOVQRYPLQRLRYVGTQLFEDNLYALVLDNGDPLNNTTPTVGASPG 154
Db 76 DIOEVQGVLLAHNOVQRYPLQRLRYVGTQLFEDNLYALVLDNGDPLNNTTPTVGASPG 135
QY 155 GLRELQTLSTLEILKGLVLIORNPOLCYODTILMKDIFHKNNQALALTIDINRSRACHPG 214
Db 136 GLRELQTLSTLEILKGLVLIORNPOLCYODTILMKDIFHKNNQALALTIDINRSRACHPG 195
QY 215 SPMCKSGKMGESSEDCOSLRTVYAGGACARCKGPLPTDCHEDQAACTGPKHSDDLAC 274
Db 196 SPMCKSGKMGESSEDCOSLRTVYAGGACARCKGPLPTDCHEDQAACTGPKHSDDLAC 255
QY 275 LHFHNSGICELHCPALVYVNTDFTESMNPREGRTTFGASCTYACRYNLTSDVGS----- 329
Db 256 LHFHNSGICELHCPALVYVNTDFTESMNPREGRTTFGASCTYACRYNLTSDVGSCTLYVC 315

```

```

QY 330 ----- 329
Db 316 PLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREVRANTSANIQEFAGCKKIFGSLA 375
QY 330 ----- 329
Db 376 FLPSFDGDPASNTAPLOPQLOVFEILEITGYLTISAMPDLSPLDSVFQNLQVIRGRI 435
QY 330 ----- 329
Db 436 LHNAGYSLTLOGLGI SWGLRSLRELGSGLALIHNNTHLCFVHTVPMQDLERNPQALLH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACIHLQCARGHCMRGPTQCVNCSQFLRGQCEVECRVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPCQPONGSVTCGPEADQCVACAHYKDPFCVARKPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 PPDEGACQPCPINCTHSCVDLDKCPAEQASPLSYSAVVGILLVVLGVVFGILI 675
QY 330 ----- 329
Db 676 KRROOKIKYTRRLQETELVEPLTPSGAMPNOAQRILKTELKRYKVLGSGAFGIVY 735
QY 330 ----- 329
Db 736 KGIWIPQGENYKIPVAKVLENTSPKANKETILDEAYVMAGVSPYVRLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTLMPYCLLDHYRENRGLSGODLLMCMQIAKMSYLEDVRLVHRDLAARVYLVKS 855
QY 330 ----- 329
Db 856 PNHYKIDFGLARLLDIDETEHADGCKVPIKMALESILRRRTFHOSDVSYGTWVEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYIMVCMWIDSECRPRELVE 975
QY 330 ----- 329
Db 976 FSRMARDPQRFVIONEDLGPASPLDSTFYRSLLLEDDMDGLVDAEEXLVPOGFFCPDP 1035
QY 330 --GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 387
Db 1036 APGAGWHHRSSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 1095
QY 388 KGLSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPOPEYVNOQDVAPRPPSPREG 447
Db 1096 KGLSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPOPEYVNOQDVAPRPPSPREG 1155
QY 448 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEYLTPOGGAAPRPHPPAFSP 507
Db 1156 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEYLTPOGGAAPRPHPPAFSP 1215
QY 508 AFDNLXYWDODPPERGAPSTFKGPTTAENPEYLGIDVP 546
Db 1216 AFDNLXYWDODPPERGAPSTFKGPTTAENPEYLGIDVP 1254

```

## RESULT 8

US-09-811-115-3

```

: Sequence 3, Application US/09811115
: Patent No. US20020035736A1
: GENERAL INFORMATION:
: APPLICANT: Erickson, Shaon
: APPLICANT: Schwall, Ralph

```

APPLICANT: King, Kathleen  
TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL  
FILE REFERENCE: GENENT. 034A  
CURRENT APPLICATION NUMBER: US/09/811,115  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/189,844  
PRIOR FILING DATE: 2000-03-16  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 1255  
TYPE: PRN  
ORGANISM: Homo sapiens  
US-09-811-115-3

Query Match 64.5%; Score 2410.5; DB 10; Length 1255;  
Best Local Similarity 41.2%; Pred. No. 9.1e-134;

Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LARGASTOYCTGTMKRLPASPETHLDMRLHLYOGCQVVGNNLELTYLPTNASLSFLQ 94  
DB 16 LPGAASTOYCTGTMKRLPASPETHLDMRLHLYOGCQVVGNNLELTYLPTNASLSFLQ 75  
QY 95 DIOGNGYVLIANNQVYVLOLRIRYGTQLEEDNVALAVDNGDPLNNTTPTTGASPG 154  
DB 76 DIOGNGYVLIANNQVYVLOLRIRYGTQLEEDNVALAVDNGDPLNNTTPTTGASPG 135  
QY 155 GLRELQRLSTLTLKGVLLIQRNPOLCYODTIIMKDIFFHKNOLATLIDITNRSRACHPC 214  
DB 136 GLRELQRLSTLTLKGVLLIQRNPOLCYODTIIMKDIFFHKNOLATLIDITNRSRACHPC 135  
QY 215 SPKCKSGRCGSESEDCQSLTRTVYVCAAGCAGCAGPLTDCCHQCAAGCTGPKHSDCLAC 274  
DB 196 SPKCKSGRCGSESEDCQSLTRTVYVCAAGCAGCAGPLTDCCHQCAAGCTGPKHSDCLAC 255  
QY 275 LHNHNSGICGLHCPALVTYVTDTFESMPNPEGRTTGASCVTACPVYVSTDVGS----- 329  
DB 256 LHNHNSGICGLHCPALVTYVTDTFESMPNPEGRTTGASCVTACPVYVSTDVGSCTLVC 315  
QY 330 ----- 329  
DB 316 PLHNGEYTAEDGTORCEKSKPCARVYCYGLMEHLREAVRANTSANIOEFAGCKKIGSLA 375  
QY 330 ----- 329  
DB 376 FLPESEFGDPSANTAPLOPELOVFTLEETITGYLYISAMPDSLPLDSVFONLQVIRGRI 435  
QY 330 ----- 329  
DB 436 LHNQAVSLTLQGLIGISWGLRLSRLBELGSLALIHNTHTLFCFVHTVPMDQLFRNPQALLH 495  
QY 330 ----- 329  
DB 496 TANRPEDECVSEGLACHQCLARGHCWGPPTQCVCNCSQFLRGOECVEECNVLQGLPREVY 555  
QY 330 ----- 329  
DB 556 NARHCLPCHPECOPONGSVTCFSPREADOCVACAHRKDPFCVACRPSGVKPLSYPMK 615  
QY 330 ----- 329  
DB 616 FPBEGACQCPINCTHSCVLDLDDKCPABQASPLTSIYSANVAILLVVLCVFGIL 675  
QY 330 ----- 329  
DB 676 KRQOKIRKYTRMRLLQETELVEPLTPSGAMPNQAMRLKETELRKVKVLSGAGFTVY 735  
QY 330 ----- 329  
DB 736 KGIWPDGENVKIPVAIKVLENTSPKANKEILDEAVYVAGVSPYVSRLLGICLTSTVO 795  
QY 330 ----- 329

DB 796 LVTLQMPYGLCLLDHVRNRCRLGSODLLNMCQIAKGMSTLEDVRYVHRDLAARNLVYS 855  
QY 330 ----- 329  
DB 856 PNHVKITDFGLARLDIDETEVHADGKVPKIMMALESILRRRTHQSDVYSGVTVMEL 915  
QY 330 ----- 329  
DB 916 MTEGAKPYDIPAREIPDLLEKGERLPQPICTIDVYIMVKWMIDSECRPRELVS 975  
QY 330 ----- 329  
DB 976 FSRMARDPQRFVVIQNEEDLCPASPLDSTFYRLSLEDMDKDLDAEETVLPQGFECPPD 1035  
QY 330 ----- 329  
DB 1036 APGAGVYVHHRHSSSTRSGGDLTLGLSEPEEAPRPLAPSEAGSDVFDGLGMAA 1095  
QY 388 KGLQSLPTHPSPLOQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVPRPSPREG 447  
DB 1096 KGLQSLPTHPSPLOQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVPRPSPREG 1155  
QY 448 PLPAARAGATLERAKTSLSGKNGVYVDVAFGAVENPEYLPQGAAPQPPPAFSP 507  
DB 1156 PLPAARAGATLERAKTSLSGKNGVYVDVAFGAVENPEYLPQGAAPQPPPAFSP 1215  
QY 508 AFNLYYWDODPPERGAPSTFKGTPTAENPEYLGIDVP 546  
DB 1216 AFNLYYWDODPPERGAPSTFKGTPTAENPEYLGIDVP 1254

#### RESULT 9

US-09-854-356-1

Sequence 1, Application US/09854356

Patent No. US20020177567A1

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Gneysen, Dirk

APPLICANT: Corixa Corporation

APPLICANT: SmithKline Beecham Biologicals S. A.

TITLE OF INVENTION: HER-2/neu Fusion Proteins

FILE REFERENCE: 014058-009810PC

CURRENT APPLICATION NUMBER: US/09/854,356

CURRENT FILING DATE: 2001-05-09

PRIOR APPLICATION NUMBER: US 09/493,480

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: US 60/117,976

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 26

SOFTWARE: Patentlin Ver. 2.1

SEQ ID NO 1

LENGTH: 1255

TYPE: PRN

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: human HER-2/neu protein

NAME/KEY: DOMAIN

LOCATION: (1)..(653)

OTHER INFORMATION: extracellular domain (ECD)

NAME/KEY: DOMAIN

LOCATION: (676)..(1255)

OTHER INFORMATION: Intracellular domain (ICD)

NAME/KEY: DOMAIN

LOCATION: (990)..(1255)

OTHER INFORMATION: phosphorylation domain (PD)

NAME/KEY: DOMAIN

LOCATION: (990)..(1048)

OTHER INFORMATION: fragment of the phosphorylation domain, preferred

US-09-854-356-1

Query Match 64.3%; Score 2405.5; DB 9; Length 1255;  
Best Local Similarity 41.1%; Pred. No. 1.8e-133;

Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY	35	LARAASTOVCSTGDMKRLRPASEPETHMLRHXOGCOVONGNELYLPYPMASLSLQ	94
Db	16	LPRMASTOVCSTGDMKRLRPASEPETHMLRHXOGCOVONGNELYLPYPMASLSLQ	75
QY	95	DIOEVGYVLIAHNOYQOYPLORLRIVRGTOJLEEDNYALAVLNDGDPYNNTPPVYASPG	154
Db	76	DIOEVGYVLIAHNOYQOYPLORLRIVRGTOJLEEDNYALAVLNDGDPYNNTPPVYASPG	135
QY	155	GLRELOLRSLTELLKGVLLQORNPOLCYDDTILMDDIFKNNQALLTLIDTNRSPACHPC	214
Db	136	GLRELOLRSLTELLKGVLLQORNPOLCYDDTILMDDIFKNNQALLTLIDTNRSPACHPC	195
QY	215	SPMKGSRGCGESEDQSLTRFVACAGGARGKGPLPTDCCHEGCAGCTGPKHSDCLAC	274
Db	196	SPMKGSRGCGESEDQSLTRFVACAGGARGKGPLPTDCCHEGCAGCTGPKHSDCLAC	255
QY	275	LHFHNSICELHCPALVTTYNTDTFESMPNPEGRHTYFGASCVTACRYNLSTDVGS----	329
Db	256	LHFHNSICELHCPALVTTYNTDTFESMPNPEGRHTYFGASCVTACRYNLSTDVGSCTLVC	315
QY	330	-----	329
Db	316	PLHNOEYTAEDGTORCEKCKPCARCYGGLGMEHLREYNAVYSAPIOERAGCKITFGSLA	375
QY	330	-----	329
Db	376	FLPESFDGDPASMTAPLOPEOLOVFETLEBITGYLISAMPDPLDSYFONLQVINGRI	435
QY	330	-----	329
Db	436	LHNGAYSLTLOGIGISWLGRLSRLBELGSLALIHNTHLCEFYHTVPMDOLEFRNPQALLH	495
QY	330	-----	329
Db	496	TANPEDECVEGELACHOLCARGHCWGPBTQCVCNSQFLRGQECVEGCRYLOGLPREYV	555
QY	330	-----	329
Db	556	NARHCLPCHPBCOPONGSVTCFGRPADQVACAHYKDPFCVACRPSGVKPDLSYMPIWK	615
QY	330	-----	329
Db	616	FRPEGACQOPCINCTHSCYVDLDDKCGAPABQASPLTISIISAVGILLVVLGVVFGILI	675
QY	330	-----	329
Db	676	KRROOKIRKYMYRRLQETELVEPLTPPSGAMPNOAQMRILKETELRKVKVLSGAFGVY	735
QY	330	-----	329
Db	736	KGIWIPGCENVKIPVAIKVLRENTSPKANKELIDEAYVWAGVSPVSRLLGICLTSTVQ	795
QY	330	-----	329
Db	796	LVYQLMPYGCILDHVRENRGRLSGODLMMWCQIAKMSYLEDVRLVHRLAARVLYKS	855
QY	330	-----	329
Db	856	PNHVKITDFGLARLLIDIDETEHADGKVPYIKMALESILRRRFTHQSOWSVYQVWEL	915
QY	330	-----	329
Db	916	MTFPAKYDGIAPAREIPDLEKEBERLPORPICHIDYUMLMVCWMLDSECRFRFELVSE	975
QY	330	-----	329
Db	976	FSMARDPORFVVUIONEDLRASPDLSTFYRSLJEEDDDGDLVDAEEXLVPOQGFCDP	10355
QY	330	--GAGGVVHHRRHSSSTRSGGDLTLGLRPSSEEARSPFLASBAGSDVDFGDMGMAA	387
Db	1036	APFGAGVHHRRHSSSTRSGGDLTLGLRPSSEEARSPFLASBAGSDVDFGDMGMAA	10959

QY	388	KGLOSLTHHPSPLORSEEDPTPLPSESDGVAAPLTCSPOEYVQNPQVRAQPSPREG	447
Db	1096	KGLOSLTHHPSPLQKRSSEDPPTPLPSESDGVAAPLTCSPOEYVQNPQVRAQPSPREG	1155
QY	448	PLPARPAGATLERAKTILSPKNGVYKDVAFEGAVENBEYLTPOGGAPPOHPPAFSP	507
Db	1156	PLPARPAGATLERPKTILSGKNGVYKDVAFEGAVENBEYLTPOGGAPPOHPPAFSP	1215
QY	508	AFDNLVYWDODPERGAPSPSTFKGPTAEENPEYLTGLDVP	546
Db	1216	AFDNLVYWDODPERGAPSPSTFKGPTAEENPEYLTGLDVP	1254

```

RESULT 10
US-09-930-125-2
: Sequence 2, Application US/09930125
: Publication No. US20020193328A1
: GENERAL INFORMATION:
: APPLICANT: Hand-Zimmerman, Susan
: APPLICANT: Cheever, Martin A.
: APPLICANT: Foy, Teresa M.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Kalos, Michael D.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Vedvick, Thomas S.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
: OF HER-2/NEU-ASSOCIATED MALIGNANCIES.
: FILE REFERENCE: 210121.544
: CURRENT APPLICATION NUMBER: US/09/930.125
: CURRENT FILING DATE: 2001-08-14
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 1255
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-930-125-2

```

Query Match	64.3%	Score 2405.5	-DB 9	Length 1255
Best Local Similarity	41.1%	Pred. No. 1.8e-133		
Matches 509; Conservative	0	Mismatches 3	Indels 727	Gaps 1

QY	35	LARGAASQVCTGIDMKRLRCLASPETHIDMYRLHYOGGQVQGNELTYLPTNLSLSTQ	94
Db	16	LRPGAASQVCTGIDMKRLRCLASPETHIDMYRLHYOGGQVQGNELTYLPTNLSLSTQ	75
QY	95	DIOEQGVVLLAHNOVROVPLQRLRYRGTOLEFEDNYLALVLDNDGPLNNTTPVTGASPG	154
Db	76	DIOEQGVVLLAHNOVROVPLQRLRYRGTOLEFEDNYLALVLDNDGPLNNTTPVTGASPG	133
QY	155	GLRELOLRSLTEILKGVLYLQIRNPOLCYQDPTILMKDIFPKNNQALATLIDNRSRACHPC	214
Db	136	GLRELOLRSLTEILKGVLYLQIRNPOLCYQDPTILMKDIFPKNNQALATLIDNRSRACHPC	199
QY	215	SPMCKSGHCWGESSESDCOSLTRYVCAGGCACRKGPLPTDCCHEDCCAMAGCTPRKHSIDLAC	274
Db	196	SPMCKSGHCWGESSESDCOSLTRYVCAGGCACRKGPLPTDCCHEDCCAMAGCTPRKHSIDLAC	255
QY	275	LHFNHSGICELHCPALVYNTDTEFESMPNPEGRTYFGASCYTACRYNYLSTDVGS----	322
Db	256	LHFNHSGICELHCPALVYNTDTEFESMPNPEGRTYFGASCYTACRYNYLSTDVGSCTLYVC	311
QY	330	-----	322
Db	316	PLHNOEYVAEDGTQRCCKSCPKCARVCYGLGMEHLREYBAVTSANIOEFAGCKKIFGSLA	374
QY	330	-----	322
Db	376	FLPESFDGDPASNTAPLQPELOLQVETLEBITGYLYISAMPDSLPLDSLVPONLQVIRGRI	431
QY	330	-----	322

```

Db 436 LHNAGVSLTQIGISMLGSLRSLRELGSLALIHNTHTLCEVHTVPWDQLFRRPHQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEBLACHQCLARGHCMPGPTQCVNCSQFLRGQCEVECHVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECOPONGSVTCFGEPAQOCVACAHYKDPFCVARCPSGVKPDLSYPMWK 615
QY 330 ----- 329
Db 616 FPDEBACOPCPINCHSCVDLDDKGCAPABQASPLTISIISAVGILLVVLGVGEGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKYMRLRLQETELVEPLTPSGAMPNOAQMRILKETELRKVKVLSGAGCTYV 735
QY 330 ----- 329
Db 736 KGIWIDGENVKIPVAIKVLENTSPRANKELIDEAVYVAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCCLDHYENKRGSLQDLNMCQIAKMSYLEVRLVHRDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKTIDFGLARLDIDETEHADGKVPKIMMALESILRRFTHQSDVMSGYVTWEL 915
QY 330 ----- 329
Db 916 MTEGAPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVKCMTIDSECRPRFRELVS 975
QY 330 ----- 329
Db 976 FSRMADPQRFVYIQNEDLGASPLDSTFRSLLEDDMDGLVDAEYTLVPQGGFCPCPD 1035
QY 330 ----- 329
Db 1036 APGAGMHHRRSSSTRSGGDLTLGLEPSEBAPRSLAPSEGASDVFDGLMGAA 1095
QY 330 ----- 329
Db 1096 KGIQSLPTDHPSPLOKRSSEPTVPLPSETDGYAPLTCSPQRYVNOQDVRQPPSPREG 1155
QY 448 PLPAAPAGATTLERAKTLSPGKNGVYKDVAFGAVENPEYLPQGAAPQPHPPAFSP 507
Db 1156 PLPAAPAGATTLERAKTLSPGKNGVYKDVAFGAVENPEYLPQGAAPQPHPPAFSP 1215
QY 508 AFQNLTYWDDPPERGAPESTFKGPTAENPEYLGLDVP 546
Db 1216 AFQNLTYWDDPPERGAPESTFKGPTAENPEYLGLDVP 1254

```

## RESULT 11

```

US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication NO. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRF

```

ORGANISM: Homo sapiens  
US-09-441-411-6

Query Match 64.3%; Score 2405.5; DB 9; Length 1255;  
Best Local Similarity 41.1%; Pred. No. 1.8e-133;  
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

```

QY 35 LARGAATQVCTGDMKRLRASPETHLDMRLHYOGCCVVOGNLELTYLPTNASLSFLQ 94
Db 16 LPPGAATQVCTGDMKRLRASPETHLDMRLHYOGCCVVOGNLELTYLPTNASLSFLQ 75
QY 95 DIOEVQGYVLIANNQVQVLORLRYRGQLEEDNATLALVNDGPRLNNTTYTGASPG 134
Db 76 DIOEVQGYVLIANNQVQVLORLRYRGQLEEDNATLALVNDGPRLNNTTYTGASPG 135
QY 155 GLRELQRLSTELIKGVLLORNPOLCYOPTIIMKDFHKNOLATLIDITNSRACHPC 214
Db 136 GLRELQRLSTELIKGVLLORNPOLCYOPTIIMKDFHKNOLATLIDITNSRACHPC 195
QY 215 SPMCKGSRGWESSEDCQSLTRTVACAGCARGKPLPTDCCHBQCAAGCTGPRHSDCLAC 274
Db 196 SPMCKGSRGWESSEDCQSLTRTVACAGCARGKPLPTDCCHBQCAAGCTGPRHSDCLAC 255
QY 275 LHFNHSGICELHCALVTYNTDIFESMPNPEGATYTGASCVTACPYVYSTDVGS 329
Db 256 LHFNHSGICELHCALVTYNTDIFESMPNPEGATYTGASCVTACPYVYSTDVGSCTLYC 315
QY 330 ----- 329
Db 316 PLNQEVTAEADGTQRCCKPCARVCYGLGMEHLREAVRANTSANIOEFACCKIKGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLOPEQLQVETLEITGYLTISAMPDSLPLDSFQNLQVINGRI 435
QY 330 ----- 329
Db 436 LHNAGVSLTQIGISMLGSLRSLRELGSLALIHNTHTLCEVHTVPWDQLFRRPHQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEBLACHQCLARGHCMPGPTQCVNCSQFLRGQCEVECHVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECOPONGSVTCFGEPAQOCVACAHYKDPFCVARCPSGVKPDLSYPMWK 615
QY 330 ----- 329
Db 616 FPDEBACOPCPINCHSCVDLDDKGCAPABQASPLTISIISAVGILLVVLGVGEGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKYMRLRLQETELVEPLTPSGAMPNOAQMRILKETELRKVKVLSGAGCTYV 735
QY 330 ----- 329
Db 736 KGIWIDGENVKIPVAIKVLENTSPRANKELIDEAVYVAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCCLDHYENKRGSLQDLNMCQIAKMSYLEVRLVHRDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKTIDFGLARLDIDETEHADGKVPKIMMALESILRRFTHQSDVMSGYVTWEL 915
QY 330 ----- 329
Db 916 MTEGAPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVKCMTIDSECRPRFRELVS 975
QY 330 ----- 329

```

Db 976 FSRMARDPQRFVYIQLNEDLGPASPLDSTFYRSLLEDDMDGLVDVDEEYLVPOQGFCCPD 1035  
 QY 330 --GAGVWHHRHRSSTRSGGDLTLGLEPSEEPASPLASGAGSDVFDGLGAA 387  
 Db 1036 APGAGVWHHRHRSSTRSGGDLTLGLEPSEEPASPLASGAGSDVFDGLGAA 1095  
 QY 388 KGLSLPTHDSPLQRYSEDPVPLPSETDGYVAPLTCSQPEVYNQPDVROPSPREG 447  
 Db 1096 KGLSLPTHDSPLQRYSEDPVPLPSETDGYVAPLTCSQPEVYNQPDVROPSPREG 1155  
 QY 448 PLPARAGATLERAKTLSSGKNGVADVAFGAVENPEYLTQGGAAQPHPPAFSP 507  
 Db 1156 PLPARAGATLERAKTLSSGKNGVADVAFGAVENPEYLTQGGAAQPHPPAFSP 1215  
 QY 508 AFDLYWDDPPERGAPSTFKCTPAENPEYGLDVP 546  
 Db 1216 AFDLYWDDPPERGAPSTFKCTPAENPEYGLDVP 1254

RESULT 12  
 US-09-821-883-5  
 ; Sequence 5: Application US/09821883  
 ; Patent No. US20020061310A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; APPLICANT: Vidovic, Damir  
 ; APPLICANT: Gradovic, Thomas  
 ; TITLE OF INVENTION: Compositions and Methods for Dendritic  
 ; FILE REFERENCE: 7636-0022.30  
 ; CURRENT APPLICATION NUMBER: US/09/821,883  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: US 60/193,504  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 479  
 ; TYPE: PR  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HER300\*-rGM-CSF construct  
 US-09-821-883-5

Query Match 58.5%; Score 2188.5; DB 10; Length 479;  
 Best Local Similarity 61.7%; Pred. No. 3.3e-121; Indels 211; Gaps 4;  
 Matches 426; Conservative 22; Mismatches 31;

QY 1 MRAAPLLARASLSLGFLLFPLFMDRSYLAKELARGAASVQVCTGTDKMLRPASPET 60  
 Db 1 MRAAPLLARASLSLGFLLFPLFMDRSYLAKELARGAASVQVCTGTDKMLRPASPET 60  
 QY 61 HLDMLRLHYGCGVQVGNLELTYLPTNASLSFLDDIOEVGYVLIANNOVROVPLQRLRI 120  
 Db 61 HLDMLRLHYGCGVQVGNLELTYLPTNASLSFLDDIOEVGYVLIANNOVROVPLQRLRI 120  
 QY 121 VRGTQLEEDNVYALAVLNDGPDLPNTVTGASPGELRELOLRSTELLKGGVLIQRNPOL 180  
 Db 121 VRGTQLEEDNVYALAVLNDGPDLPNTVTGASPGELRELOLRSTELLKGGVLIQRNPOL 180  
 QY 181 CYODTILMKDIFHKNNQALTLTIDTNRSRACHPCSPMKCSRCWGESSEDCOSLTRIVCA 240  
 Db 181 CYODTILMKDIFHKNNQALTLTIDTNRSRACHPCSPMKCSRCWGESSEDCOSLTRIVCA 240  
 QY 241 GGCARCGRLPTDCCHQCAAGCTGPKHSDCLACLHNHSGICGLCPALVTYTTDFES 300  
 Db 241 GGCARCGRLPTDCCHQCAAGCTGPKHSDCLACLHNHSGICGLCPALVTYTTDFES 300  
 QY 301 MPNEGRYTGASCVTPCNVYLTSDVSGAGVWHHRHRSSTRSGGDLTLGLEPSEE 360  
 Db 301 MPNEGRYTGASCVTPCNVYLTSDVSGAGVWHHRHRSSTRSGGDLTLGLEPSEE 360  
 QY 361 EAPRSPLAPSEAGSDVFDGLGAAKGLSLPTHDSPLQRYSEDPVPLPSETDGYV 420

Db 332 ----- 331  
 QY 421 APLTCSQPEVYNQPDVROPSPREGRLPARAGATLERAKTLSSGKNGVADVAFAG 480  
 Db 332 ----- 334  
 QY 481 GAVENPEYLTQGGAAQPHPPAFSPADNLYWDDPPERGAPSTFKCTPAENPEY 540  
 Db 335 ----- 338  
 QY 541 LGLDVPAARASPSSTOPMEHYNALQEARLLNLSRDTAENETVEYISEMFDLOEP 600  
 Db 339 ----- 391  
 QY 601 TCIOTRLKLYKQGRSLTLYKGLTMMASHYKONCPPTBETSCATQIITFESKELKD 660  
 Db 392 TCYOTRLKLYKQGRSLTLYKGLTMMASHYKONCPPTBETDELEVTTFEDEFKMLKG 451  
 QY 661 FLVYIPDCWEPVOEGAPPPAAHHNNH 690  
 Db 452 FLVYIPDCWEPVOEGAPPPAAHHNNH 479

RESULT 13  
 US-09-854-356-14  
 ; Sequence 14: Application US/09854356  
 ; Patent No. US20020177567A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cheever, Martin A.  
 ; APPLICANT: Gheysen, Dirk  
 ; APPLICANT: Corixa Corporation  
 ; APPLICANT: SmithKline Beecham Biologicals S. A.  
 ; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
 ; FILE REFERENCE: 014058-009810PC  
 ; CURRENT APPLICATION NUMBER: US/09/854,356  
 ; PRIOR FILING DATE: 2001-05-09  
 ; PRIOR APPLICATION NUMBER: US 09/493,480  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: US 60/117,976  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 1256  
 ; TYPE: PR  
 ; ORGANISM: Mus sp.  
 ; FEATURE:  
 ; OTHER INFORMATION: mouse HER-2/neu protein  
 US-09-854-356-14

Query Match 52.4%; Score 1961; DB 9; Length 1256;  
 Best Local Similarity 54.3%; Pred. No. 2.4e-107; Indels 740; Gaps 3;  
 Matches 434; Conservative 23; Mismatches 67;

QY 11 AASLSLGFLLFPLFMDRSYLAKELARGAASVQVCTGTDKMLRPASPTHMLRLHYQ 70  
 Db 4 AAMCRMGFLAL-----LSPGAAGTQVCTGTDKMLRPASPTHMLRLHYQ 51  
 QY 71 GCOVQGNLELTYLPTNASLSFLDDIOEVGYVLIANNOVROVPLQRLRIYRGTQLEEDN 130  
 Db 52 GCOVQGNLELTYLPTNASLSFLDDIOEVGYVLIANNOVROVPLQRLRIYRGTQLEEDN 111  
 QY 131 YALAVIDNDGPDLPNTVTGASPGELRELOLRSTELLKGGVLIQRNPOLCYODTILMK 169  
 Db 112 YALAVIDNDGPDLPNTVTGASPGELRELOLRSTELLKGGVLIQRNPOLCYODTILMK 171  
 QY 190 DIFHKNNQALTLTIDTNRSRACHPCSPMKCSRCWGESSEDCOSLTRIVCAAGCARCGP 249  
 Db 172 DVLKNNQALPVMMDTNRSRACHPCAPCTCKDNMGWSPEDCQILTGTICTSGCARCKGR 231  
 QY 250 LPTDCHQCAAGCTGPKHSDCLACLHNHSGICGLCPALVTYTTDFESMPNEGRY 309  
 Db 250 LPTDCHQCAAGCTGPKHSDCLACLHNHSGICGLCPALVTYTTDFESMPNEGRY 309





```

QY 330 ----- 329
Db 770 KEIDEAVYMGVSGPYVSRLLGICLTSTVOLVQLMPYGCGLDHYREHGRGLSSQDLN 829
QY 330 ----- 329
Db 830 MCVQIAKMSYLELVRLVHRDLAARNLVKSPNHVKITDFGLARLLDIDETEHADGKGV 889
QY 330 ----- 329
Db 890 PIKMALESLRRRFTHSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPQ 949
QY 330 ----- 329
Db 950 PICTIDVYIMVWKCMIDSECRPRRELVSERKARDPQRFVYIQNEDLGSSPMDSF 1009
QY 330 ----- 356
Db 1010 YRSLLEDMDGDLVDAEELVPOQGFSPDPTPGTSTAHRHRSSSTRSGGELTGLE 1069
QY 357 PSEEAAPSPAPSPGASDVFDGDLGGAAGKGLDGLPTHPSPLOQRTSEDPVLPSET 416
Db 1070 PSEEBPPSPAPSPGASDVFDGDLGAGVTKGLDGLSPHDLSPLOQRTSEDPVLPSET 1129
QY 417 DGVAFLCSPQPEYVNPQVPPSPREGPLPAPAPAGATLERAKTILSPCKNGVQV 476
Db 1130 DGVAFLCSPQPEYVNPQVPPSPREGPLPAPAPAGATLERAKTILSPCKNGVQV 1189
QY 477 FAFGAVENPEYLPOGGAQAPHPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAE 536
Db 1190 FAFGAVENPEYLREGTASPPHSPAFSPAFDNLVYWDONSSSQGPPSPNFEETPTAE 1249
QY 537 NPEYGLDVP 546
Db 1250 NPEYGLDVP 1259

```

## RESULT 15

```

US-09-854-356-2
Sequence 2, Application US/09854356
Patent No. US2002017367A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Gheysen, Dirk
APPLICANT: Corixa Corporation
APPLICANT: SmithKline Beecham Biologicals S. A.
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
PRIORITY FILING DATE: 2001-05-09
PRIORITY APPLICATION NUMBER: US 09/493,480
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: US 60/117,976
PRIORITY FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1256
TYPE: PRT
ORGANISM: Rattus sp.
FEATURE:
OTHER INFORMATION: rat HER-2/neu protein
NAME/KEY: DOMAIN
LOCATION: (1)..(654)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (677)..(1256)
OTHER INFORMATION: intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (721)..(998)
OTHER INFORMATION: kinase domain (KD)
NAME/KEY: DOMAIN

```

```

LOCATION: (991)..(1256)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (991)..(1049)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD)
US-09-854-356-2

```

```

Query Match 52.1% Score 1948: DB 9: Length 1256;
Best Local Similarity 34.3% Pred. No. 14e-166;
Matches 433: Conservative 21: Mismatches 70: Indels 740: Gaps 3:

```

```

QY 11 AASLIGFLFLFEMLDRLSLAKELARGASTOYCTGDMKRLYPASPEHLMRLHLYQ 70
Db 4 AAMCKMGFLAL-----LPPGIAGTQVCTGDMKRLYPASPEHLMRLHLYQ 51
QY 71 GCQVVOGNLELYLPTNASLSFLDIOEVQGYLVIANQVQVPLQRLIRYGTQLEFDN 130
Db 52 GCQVVOGNLELYLVANASLSFLDIOEVQGYLVIANQVQVPLQRLIRYGTQLEFDK 111
QY 131 YALAVLDGDDPLNNTPTV-GASPGLEQLRLSTETILKGVILQRRPOLCYQDTLMK 189
Db 112 YALAVLDNRDQDVAASTPERTPEGLRELQRLSTETILKGVILKGNPOLCYQDMVLMK 171
QY 190 DIFHKNQALATLIDTNRSRACHPCSPMKGSRGSESEDCOSLTPTVACAGCARKGP 249
Db 172 DIFRKNQALAVDIDTNRSRACHPCSPMKGSRGSESEDCOSLTPTVACAGCARKGP 231
QY 250 LPTDCHEQCAAGCTGPRHSDCLACLFHNSGICELHCPALVYNTDTFESMPNPEGRT 309
Db 232 LPTDCHEQCAAGCTGPRHSDCLACLFHNSGICELHCPALVYNTDTFESMHNDEGRYT 291
QY 310 FGASCVTACPNYVLTSDVGS----- 329
Db 292 FGASCVTTCPTNYVLTSEVGSCTLVCPNNQEVTAEDTQRCCKSPCARVCYGLGMEHL 351
QY 330 ----- 329
Db 352 RGAARITSDNVOEFDCCKKIFGSLAFLPESFDGDPSSGIAPLRPQQLVFETLEITGYL 411
QY 330 ----- 329
Db 412 YISAMPDSLRLSVFQNLRIIRGRLIHDGAYSLTLOGIGHSLSRLSRELGSGLALYH 471
QY 330 ----- 329
Db 472 NAHLCFVHTVMDOLFRRPHQALHSGNRPEBDCGLEVCNSLCAHGCWGPPTQCVN 531
QY 330 ----- 329
Db 532 CSHFLRGQCEVBECHVMKGLPREVYSKRCLPCHPEQCPONSSETCFGBADQCAACAHY 591
QY 330 ----- 329
Db 592 KDSSCVAPRCPSGVKPDLSYMPIMKYPDEEGICOPCINCTHSCVDLDERGCPAQORASP 651
QY 330 ----- 329
Db 652 VTIIATVEGYLLFLILVYVVGILIKRRRQKIRKTYMRLLQETELVPLPSPGAMPQA 711
QY 330 ----- 329
Db 712 QMRILKETELRKVKVLSGAGFYVKGIMIPDGENVKIPVAIKVLENTSPRANKELDE 771
QY 330 ----- 329
Db 772 AYVAVGSPYVSRLLGICLTSTVOLVQLMPYGCGLDHYREHGRGLSSQDLNMCVQIA 831
QY 330 ----- 329
Db 832 KGMSTLELVRLVHRDLAARNLVKSPNHVKITDFGLARLLDIDETEHADGKGVPIKMA 891
QY 330 ----- 329

```

```

Db 892 LESILRRRTHQSDVMSYGVWELMTFGAKPYDGIIPAREIPDLLEKGERLPOPPICITID 951
QY 330 ----- 329
Db 952 VYIMIVKCMIDSECRPRRRELYSEFSRMARDPQRFVVIQNEDLGSPSPMDSTFYRSLE 1011
QY 330 -----GACGMVHHRRHSSSTRSGGDLTLGLEPSEEA 362
Db 1012 DDDMGDLVDAEEYLVPOGFFSPDPPTPGTSTAHRHRHSSSTRSGGELTLGLEPSEEGP 1071
QY 363 PRSPLAPSEAGSDVFPDGLGMAKGLQSLPTPHDPSPLQRYSEDPTVPLPSETDGYVAP 422
Db 1072 PRSPLAPSEAGSDVFPDGLGMAKGLQSLPHDLSPLQRYSEDPTVPLPETDGYVAP 1131
QY 423 LTCSPQPEYVNOPDVRPQPPSPREGPLPAARPAAGATLERAKTLPCKNGVVKDVFAGCA 482
Db 1132 LACSPQPEYVNOSEVQPPPLTPEGPLPVYRPAAGATLERPKTLPCKNGVVKDVFAGCA 1191
QY 483 VENPEYLTPOGGAAPQHPPPAFSPAFDNLXYWDODPPERGAPPSTFKGPTAENPEYLG 542
Db 1192 VENPEYLVPREGTASPPHPSPAFSPAFDNLXYWDQNSSEOGPPSPNEGTPTAENPEYLG 1251
QY 543 LDVP 546
Db 1252 LDVP 1255

```

Search completed: April 28, 2003, 13:44:03  
 Job time : 40.3846 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:39:54 ; Search time 20.1923 Seconds  
(Without alignments)  
1005.423 Million cell updates/sec

Title: US-09-821-883-2  
Perfect score: 3739  
Sequence: 1 MRAAPLLIARASTISLGLF.....EPVOEGAPPPAAHHHHH 690

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	64.3	1255	1 US-08-467-083-68	Sequence 68, Appl
2	2405.5	64.3	1255	1 US-08-414-417B-68	Sequence 68, Appl
3	2405.5	64.3	1255	2 US-08-486-348A-68	Sequence 68, Appl
4	2405.5	64.3	1255	2 US-08-625-101-2	Sequence 2, Appl
5	2405.5	64.3	1255	2 US-08-468-545B-68	Sequence 68, Appl
6	2405.5	64.3	1255	2 US-08-356-786-2	Sequence 2, Appl
7	2405.5	64.3	1255	3 US-08-466-680B-68	Sequence 68, Appl
8	2385.5	63.8	1255	2 US-08-484-438-8	Sequence 8, Appl
9	2224	59.5	782	3 US-09-146-283-4	Sequence 4, Appl
10	2224	59.5	782	3 US-08-579-823A-4	Sequence 4, Appl
11	2224	59.5	782	4 US-09-344-195-4	Sequence 2, Appl
12	1607	43.0	419	4 US-09-630-155-2	Sequence 2, Appl
13	1597	42.7	624	3 US-08-422-108-1	Sequence 1, Appl
14	1597	42.7	624	4 US-08-422-734-1	Sequence 69, Appl
15	1183	31.6	580	1 US-08-414-417B-69	Sequence 69, Appl
16	1183	31.6	580	2 US-08-486-348A-69	Sequence 69, Appl
17	1183	31.6	580	2 US-08-468-545B-69	Sequence 69, Appl
18	1183	31.6	580	3 US-08-466-680B-69	Sequence 69, Appl
19	761	20.4	515	2 US-09-146-283-2	Sequence 2, Appl
20	761	20.4	515	3 US-08-579-823A-2	Sequence 2, Appl
21	721	19.3	644	1 US-09-344-195-2	Sequence 2, Appl
22	721	19.3	644	1 US-08-336-708A-9	Sequence 9, Appl
23	721	19.3	911	2 US-08-484-438-10	Sequence 10, Appl
24	721	19.3	1058	2 US-08-484-438-4	Sequence 7, Appl
25	721	19.3	1210	2 US-08-484-438-7	Sequence 4, Appl
26	721	19.3	1210	2 US-08-475-035-4	Sequence 2, Appl
27	721	19.3	1308	2 US-08-484-438-2	Sequence 2, Appl

28	704.5	18.8	478	4 US-09-570-454-2	Sequence 2, Appl
29	698.5	18.7	1343	6 5183884-4	Patent No. 5183884
30	697	18.6	1342	1 US-07-978-895-4	Sequence 4, Appl
31	697	18.6	1342	2 US-08-484-438-9	Sequence 9, Appl
32	697	18.6	1342	2 US-08-473-119-4	Sequence 4, Appl
33	697	18.6	1342	2 US-08-475-352-4	Sequence 4, Appl
34	693	18.5	274	3 US-08-469-318-144	Sequence 144, App
35	693	18.5	274	3 US-08-468-609A-144	Sequence 144, App
36	693	18.5	274	4 US-08-446-872A-144	Sequence 144, App
37	693	18.5	274	4 US-08-762-227A-144	Sequence 144, App
38	693	18.5	274	5 PCT-US93-01185-144	Sequence 144, App
39	680.5	18.2	301	3 US-08-469-318-142	Sequence 142, App
40	680.5	18.2	301	3 US-08-468-609A-142	Sequence 142, App
41	680.5	18.2	301	4 US-08-446-872A-142	Sequence 142, App
42	680.5	18.2	301	4 US-08-762-227A-142	Sequence 142, App
43	680.5	18.2	301	5 PCT-US93-01185-142	Sequence 142, App
44	679	18.2	165	1 US-08-318-193-8	Sequence 8, Appl
45	678	18.1	168	1 US-08-318-193-10	Sequence 10, Appl

#### ALIGNMENTS

RESULT 1  
US-08-467-083-68  
Sequence 68, Application US/08467083  
Patent No. 5726023  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,083  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/414,417  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
TELEX: 3723836 SEDANBERRY  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-467-083-68  
Query Match 64.3% Score 2405.5; DB 1; Length 1255;  
Best Local Similarity 41.1% Pred. No. 4.8e-179;  
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;  
35 LARGAASOVCTGTDMLRLPASPETHLDMLRLHYOGGVGVGNLELYLPTNASTLSFLQ 94

```

Db 16 LPPGAATGYCTGDMKRLRPASBETHLDMRLHYGCGVVOGNLELTYLPTNASLSFLQ 75
QY 95 DIEVOGCVYLIANNOVQVLOLRIRYRGTOLEFEDNALAVLNDGDLNNTTPTGASPG 154
Db 76 DIEVOGCVYLIANNOVQVLOLRIRYRGTOLEFEDNALAVLNDGDLNNTTPTGASPG 135
QY 155 GLRELQRLSLTEILKGVLQORNPOLCYODTILMKDIFHKNNOLATLIDITNSRACHPC 214
Db 136 GLRELQRLSLTEILKGVLQORNPOLCYODTILMKDIFHKNNOLATLIDITNSRACHPC 195
QY 215 SPMCKGRGCGESSEDCQSLRTYVCAAGCARGKGPLPTDCHBQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKGRGCGESSEDCQSLRTYVCAAGCARGKGPLPTDCHBQCAAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDFTESMPNPEGRTFGASCVTACPYNYLSTDVGS 329
Db 256 LHFHNSGICELHCPALVTYNTDFTESMPNPEGRTFGASCVTACPYNYLSTDVGSCTLYC 315
QY 330 329
Db 316 PLHNOEVTAEADGTQRCCKSPCARVCGIGMEHLREAVTSANIQEFAGCKKIFGSLA 375
QY 330 329
Db 376 FLPESEFGDPAANTAPLQPEOLOVFTLEBITGYLYISAMPDLSPLSVFONLOVINGRI 435
QY 330 329
Db 436 LHNQAVSLTLQGLISWGLRSLRELGSGLALHNTHLGCVHTVPMWDLFRNPHQALH 495
QY 330 329
Db 496 TANRPEDECVGEGLACHQLCARHCWGPPTQVCNCSQFLRGQECVECHVLOGLPREVY 555
QY 330 329
Db 556 NARHCLPHECOPQNSVTCFGEADQVCAHAHYKDPFCVARGPBGVYKPDLSYMPIMK 615
QY 330 329
Db 616 FPBEGACOPCPINCHTSCVDLDDKGRPAQRASPLSIISAVVGLLVVGLVGFILI 675
QY 330 329
Db 676 KRROQKIRKTYMRLLOETELVEPLTPSGAMPQAOHRIKETELRKVYKVGSGAFSTY 735
QY 330 329
Db 736 KGIWIPDGENVKIPVAIKULRENTSPRANKELIDEAVYMAVGSPYVSRLLGICLSTVQ 795
QY 330 329
Db 796 LVTQLMPYGLDLHVENRGRGLSQDILLNMCQIAKMSYLEDRVLYHRDLAARNVLKYS 855
QY 330 329
Db 856 PNHYKITDFGLARLLDIDETEHADGKVPKIMMALESILRRRTHOSDVMSYGVYMWEL 915
QY 330 329
Db 916 MTRGAKPYDGPARELPDLLEKGBRLPPICITIDVYIMVCKMIDSECRPRELAVSE 975
QY 330 329
Db 976 FSHMARPORFVVIQNEDELGPASPLDSTFYRSLLEDDMDDIVDAEYLYVPOGFFPCPD 1035
QY 330 387
Db 1036 APAGCGVYHHRHSSSTRSGGDLTLGLEPSEEBAPRSLAPSEGASDVFDGLGAA 1095
QY 388 KGIQSLPTHPSPLOVRSDEPTVPLPSETDGYVAPLCSQPEYVNPDPVRPQPPSPREG 447

```

```

Db 1096 KGIQSLPTHPSPLOVRSDEPTVPLPSETDGYVAPLCSQPEYVNPDPVRPQPPSPREG 1155
QY 448 PLPARAGATLERAKTILSPGKNVGVDFAFGAVENPEYLTPOGGAQDPHPPPAFSP 507
Db 1156 PLPARAGATLERAKTILSPGKNVGVDFAFGAVENPEYLTPOGGAQDPHPPPAFSP 1215
QY 508 AFDNLYWDDPPPERGAPSTFKGTPTAENPEYLGIDVP 546
Db 1216 AFDNLYWDDPPPERGAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 2
US-08-414-417B-68
; Sequence 68, Application US/08414417B
; Patent No. 5801005
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESS: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELEPHONE: (206) 622-4900
; TELECOMMUNICATION INFORMATION:
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-414-417B-68

Query Match 64.3%; Score 2405.5; DB 1; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

```

```

|||||
Db 256 LHFHNSGICELHCPALVYNTDFTESMPNPGRYTFGASCVTACPYNTLSDVGSCTLYVC 315
Qy 330 ----- 329
Db 316 PLHNOEVAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIOBFACKKIFGSLA 375
Qy 330 ----- 329
Db 376 FLPSFDGDPASNTAPLOPELOVFEETLEITGYLYISAMPDSLPLDSVFONLOYIRGRI 435
Qy 330 ----- 329
Db 436 LHNCAVSLTLQGLISWLGRLSRELGSGLALIHNNHLCFVHTVPMQDLFRNPQALLH 495
Qy 330 ----- 329
Db 496 TANRPEDECVGEGLACHQLCARCHCWBPYQVCNCSQFLRGQBCVECRVLQGLPREYV 555
Qy 330 ----- 329
Db 556 NARHCLPCHPCOPONGSVTCGPEADOCVACAHYKOPPEVCANCPSPGVKPDLSMPYWK 615
Qy 330 ----- 329
Db 616 FPDEGACQPCPINCTHSCVDLDDKCPAQORASPLTISIISAVVIGLLVVLGVVIGILI 675
Qy 330 ----- 329
Db 676 KRROOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETELRKVKVLSGAFGTIVY 735
Qy 330 ----- 329
Db 736 KGIWIPGQENKIVAIKVLRENTSPRANKRILDEAVYMAVGSPYSRLGICILTSTVO 795
Qy 330 ----- 329
Db 796 LVTOLMPYCLLDHYRENRGRIGSODLLMCMQIAKMSYLEVDVRLVHRDLAARNVLYKS 855
Qy 330 ----- 329
Db 856 PNHVKITDFGLARLLIDETEHADGKVPKIMMALESILRRRPTHOSDWSYGVYVWL 915
Qy 330 ----- 329
Db 916 MTFGAKPYDGTIPAREIPDLLEKGERLPQPICTIDVYIMVWKMIIDSECRPRRELAYSE 975
Qy 330 ----- 329
Db 976 FSRMARDFQRFVIONEDLGPASPLDSTFYRSLLLEDDMDGLVDAAEYLVPOQGFPCDP 1035
Qy 330 ----- 329
Db 1036 ARGAGGVNHRHRRSSSTSGGGDLTLGLEPSEEEAPRSPPLAPSBGAGSDVVDGDLGMAA 1095
Qy 388 KGLSLPTHDSPLORYSEDPVPLPSETDGVYVAPLTCSPQREYVNOVDVPRPPSPREG 447
Db 1096 KGLDLSLPTHDSPLORYSEDPVPLPSETDGVYVAPLTCSPQREYVNOVDVPRPPSPREG 1155
Qy 448 PLPARPAGATLERAKTUSPGKNGVYKDVFAFGAVENPEYLTPOGAAPQPHPRPASP 507
Db 1156 PLPARPAGATLERAKTUSPGKNGVYKDVFAFGAVENPEYLTPOGAAPQPHPRPASP 1215
Qy 508 AFDNLYYWDODPRERGAPESTFKGPTAENPEYLGIDVP 546
Db 1216 AFDNLYYWDODPRERGAPESTFKGPTAENPEYLGIDVP 1254

```

RESULT 3  
 US-08-486-348A-68  
 : Sequence 68: Application US/08486348A  
 : Patent No. 5846538  
 : GENERAL INFORMATION:

```

APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELEPHONE: (206) 682-4800
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

```

```

Query Match 64.3%; Score 2405.5; DB 2; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;
Qy 35 LARGAASVQVCTGDMKRLPASPEETHLDMRLHLYOGCOVVOGNIETLYLPTNLSLSTLQ 94
Db 16 LPPGAASVQVCTGDMKRLPASPEETHLDMRLHLYOGCOVVOGNIETLYLPTNLSLSTLQ 75
Qy 95 DIOEVQGVYLAHNOVROVPLORLIRYVGTOLFEEDNVALAVLNDGDPPLNNTPTVPGASPG 154
Db 76 DIOEVQGVYLAHNOVROVPLORLIRYVGTOLFEEDNVALAVLNDGDPPLNNTPTVPGASPG 135
Qy 155 GLRELQLRSLTEILKGVLIQIORNPOLCYQDTILMKDIFHKNNQALATLIDTNRSRACHPC 214
Db 136 GLRELQLRSLTEILKGVLIQIORNPOLCYQDTILMKDIFHKNNQALATLIDTNRSRACHPC 195
Qy 215 SPMCKGSRKWESESDCOSLTRTYCAGCAGCARKGRLPTDCHEGCAAGCTGPKHSIDLAC 274
Db 196 SPMCKGSRKWESESDCOSLTRTYCAGCAGCARKGRLPTDCHEGCAAGCTGPKHSIDLAC 295
Qy 275 LHFHNSGICELHCPALVYNTDFTESMPNPGRYTFGASCVTACPYNTLSDVGSCTLYVC 329
Db 256 LHFHNSGICELHCPALVYNTDFTESMPNPGRYTFGASCVTACPYNTLSDVGSCTLYVC 315
Qy 330 ----- 329
Db 316 PLHNOEVAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIOBFACKKIFGSLA 375
Qy 330 ----- 329
Db 376 FLPSFDGDPASNTAPLOPELOVFEETLEITGYLYISAMPDSLPLDSVFONLOYIRGRI 435
Qy 330 ----- 329
Db 436 LHNCAVSLTLQGLISWLGRLSRELGSGLALIHNNHLCFVHTVPMQDLFRNPQALLH 495
Qy 330 ----- 329

```

```

Db 496 TANRDECEVGEGLACHQIARCHCWGPGPTQCVCNSQFLRGQCEVCECVLQGLPREV 555
QY 330 ----- 329
Db 556 NARHCLPHECOPQNGSVTCFGEADQCAAHYKDPFCVARGPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEGACOPCPINCHSCVLDLDKGCAPARASPLTISAVVGLLVVGLVGFGLI 675
QY 330 ----- 329
Db 676 KRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETELRKVKVJGSGAGCTVY 735
QY 330 ----- 329
Db 736 KGIWPDGENVKIPVAIKVIRENTSPKANKEILDEAVYVAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQAMPYGLLDHRENRGRGLSODLLNMCQIAKMSYLEDVRLVHRDLAARNVLKVS 855
QY 330 ----- 329
Db 856 PNHYKITDFGLARLLDIDETEHADGSKVPDKWMALESILRRRFTHOSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTEGAKYDGIIPAREIDLEKEGRLLPQPICTIDVYMIWKCWIMIDSECRPRELVE 975
QY 330 ----- 329
Db 976 FSRMARPORFVVYQNDLSPASPLDSTFYRLSLEDDEMDLVDAEEYLVPOQGFCCPD 1035
QY 330 ----- 387
Db 1036 APGGAGVHRRHSSSTRSGGDLTGLPSEEBAPRSLAPSGASDVFDGGLGAA 1095
QY 388 KGIQSLTHTDPSLQRYSEDPTVPLPSETDGYVAPLTCSPQREXVNOVDVRRPQPSREG 447
Db 1096 KGIQSLTHTDPSLQRYSEDPTVPLPSETDGYVAPLTCSPQREXVNOVDVRRPQPSREG 1155
QY 448 PLPARAGATLERAKTLSEKNGVAVDAFAGAVENPEYLPQGGAAQPPHPPAFSP 507
Db 1156 PLPARAGATLERAKTLSEKNGVAVDAFAGAVENPEYLPQGGAAQPPHPPAFSP 1215
QY 508 AFDLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 546
Db 1216 AFDLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 1254

```

```

RESULT 4
US-08-625-101-2
; Sequence 2, Application US/08625101
; Patent No. 5869445
;
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-Apr-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-625-101-2

```

```

Query Match 64.3%; Score 2405.5; DB 2; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGASTQVCTGDMKRLPASPETHLDMLRHLVQCGVVOGNTLELYPTNALSFLQ 94
Db 16 LPGAATQVCTGDMKRLPASPETHLDMLRHLVQCGVVOGNTLELYPTNALSFLQ 75
QY 95 DIOEVQGVLIANQVQVPLQRLIRYGTQLEFEDNALAVLNDGDPILNNTPTTGASPG 154
Db 76 DIOEVQGVLIANQVQVPLQRLIRYGTQLEFEDNALAVLNDGDPILNNTPTTGASPG 135
QY 155 GLRELOLRSLLETKGVLIOBNPOLCYODTILMKDIFHKNNOLATLITNRSRACHPC 214
Db 136 GLRELOLRSLLETKGVLIOBNPOLCYODTILMKDIFHKNNOLATLITNRSRACHPC 195
QY 215 SPMKSGRCWGESESDQSLTRIVYAGAGCARCKGPLPTDCHEGCAAGCTGPKHSDCLAC 274
Db 196 SPMKSGRCWGESESDQSLTRIVYAGAGCARCKGPLPTDCHEGCAAGCTGPKHSDCLAC 255
QY 275 LHFHSGICELHCPALVTYMTDFESMPNPEGRTFGASCVTACRYVLTSDGCS 329
Db 256 LHFHSGICELHCPALVTYMTDFESMPNPEGRTFGASCVTACRYVLTSDGCS 315
QY 330 ----- 329
Db 316 PLHNGEYTAEDGTORCEKSKPCARVCYGLGMEHLREAVTANSIOEFAGCKKIEGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLOPOLQVETLEETITGYLIYISAMPDSLPLDSVFNQLYINGRI 435
QY 330 ----- 329
Db 436 LHNQAVSLTQGLGISWGLRSLRELQSGALIHNTHTLCFVHTVPWDQLFRNPQOALLH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHQIARCHCWGPGPTQCVCNSQFLRGQCEVCECVLQGLPREV 555
QY 330 ----- 329
Db 556 NARHCLPHECOPQNGSVTCFGEADQCAAHYKDPFCVARGPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEGACOPCPINCHSCVLDLDKGCAPARASPLTISAVVGLLVVGLVGFGLI 675
QY 330 ----- 329
Db 676 KRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETELRKVKVJGSGAGCTVY 735

```

```

QY 330 ----- 329
Db 736 KGIWIPGDNKIPVAIKVLRENTSPKANKRIIDEAIVMAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTLMPYGLLDHVRENRLGSQDLLNMCQIAKMSYLEVRLVHRLAARNLVKS 855
QY 330 ----- 329
Db 856 PNHVKTDFGLARLLDIDETEFYHADGKVPKIMMALESILRRPTHOSDVSYGVYWEI 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVCKMIDSECRPRELVS 975
QY 330 ----- 329
Db 976 FSRMARPOREVIQNEIDLPASPLDSTFYRSLLEDDMDGLVDAEYLVPOGFFCPDP 1035
QY 330 ----- 387
Db 1036 APGAGVHHRRSSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVVDGOLGMAA 1095
QY 388 KGLQSLPTHDSPLQKRSDEPTVPLPSETDGYVAPLTCSPQPEYVNOVDVPPSPREG 447
Db 1096 KGLQSLPTHDSPLQKRSDEPTVPLPSETDGYVAPLTCSPQPEYVNOVDVPPSPREG 1155
QY 448 PLPARAPGATLERAKTLSPGKNGVADVAFAGAVENPEYLTPQGAAPRPPAPSP 507
Db 1156 PLPARAPGATLERAKTLSPGKNGVADVAFAGAVENPEYLTPQGAAPRPPAPSP 1215
QY 508 AFDLVYWDODPBERGAPSPSTFKGTPTAENPEYLGDPV 546
Db 1216 AFDLVYWDODPBERGAPSPSTFKGTPTAENPEYLGDPV 1254

RESULT 5
US-08-468-545B-68
Sequence 68 Application US/08468545B
Patent No. 5876112
GENERAL INFORMATION:
APPLICANT: Cheever, Marlin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match 64.3%; Score 2405.5; DB 2; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGAATQVCTGTDMKRLPASPETHLMDLRHLVYGGVQVQGNLELYLPTNASLSFLQ 94
Db 16 LPFGAATQVCTGTDMKRLPASPETHLMDLRHLVYGGVQVQGNLELYLPTNASLSFLQ 75
QY 95 DIOEVQGVYLLAHNOVROVPLRIYVGTQVLFEDNALAVLDGDDPLNNTPTVTSAPG 154
Db 76 DIOEVQGVYLLAHNOVROVPLRIYVGTQVLFEDNALAVLDGDDPLNNTPTVTSAPG 135
QY 155 GLREIQLRSLTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQALATLIDTNRSRACHPC 214
Db 136 GLREIQLRSLTEILKGVLIQGNPOLCYQDTILMKDIFHKNNQALATLIDTNRSRACHPC 195
QY 215 SPMCKGSRKMGESSEDCSLTRTYVAGGACARCKGPLPTDCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRKMGESSEDCSLTRTYVAGGACARCKGPLPTDCHEQCAAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVYNTDFESMPNPRGRTFGASCTACPRYNLSTDVGS ---- 329
Db 256 LHFHNSGICELHCPALVYNTDFESMPNPRGRTFGASCTACPRYNLSTDVGSCTLVLC 315
QY 330 ----- 329
Db 316 PLHNOEYTAEDGTQRCCKSKPCARVCYGLGMEHLREVRATVANSIOEFAGCKKIFSSLA 375
QY 330 ----- 329
Db 376 FLPSFDDPASNAPLQPEOLQVFELEITGYLISAMPDSLPLDSVFNOLVIRGRI 435
QY 330 ----- 329
Db 436 LHNCAVSLTLOGLISWLGSLRLSRLGSLALIHNNHLCFVHYVPMQDLFRNHQALLH 495
QY 330 ----- 329
Db 496 TANRDECVGEGLACHQLCARCHQMGPRPTQVCVNSQFLRGQCEVCRIVLGGLPREV 555
QY 330 ----- 329
Db 556 NARHCLPCHPCQPNQNGSVTCGPEADQVACAHYKDPFCVAPRPSGVKPPDLSTYPIWK 615
QY 330 ----- 329
Db 616 FPDEGACQPCPINCCTHSCVDLDKCPAEORASPLTISIAVVGILLVYLVGVVFGILI 675
QY 330 ----- 329
Db 676 KRROKIRKTYMRRLQETELVEPLTPSGAMPNQAQMLKETELRKVVLGSAFGTVY 735
QY 330 ----- 329
Db 736 KGIWIPGDNKIPVAIKVLRENTSPKANKRIIDEAIVMAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTLMPYGLLDHVRENRLGSQDLLNMCQIAKMSYLEVRLVHRLAARNLVKS 855
QY 330 ----- 329
Db 856 PNHVKTDFGLARLLDIDETEFYHADGKVPKIMMALESILRRPTHOSDVSYGVYWEI 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVCKMIDSECRPRELVS 975

```

QY 330 ----- 329  
Db 976 FSHMARDQRFVVIQNEEDLGPASPLDSTFYRSLLEDDMDGLVDABEYLVPQGFECPPD 1035  
QY 330 --GAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMAA 387  
Db 1036 APAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMAA 1095  
QY 388 KGQSLPTHPSPLOKRSSEPTVPLPSEIDGYAPLTCSPQPEYVNOQDVPQPPSPREG 447  
Db 1096 KGQSLPTHPSPLOKRSSEPTVPLPSEIDGYAPLTCSPQPEYVNOQDVPQPPSPREG 1155  
QY 448 PLPAARAGATLERRAKTLSPGKNGVYKDVAFGAVENPEYLPQGGAAQPPAPAFSP 507  
Db 1156 PLPAARAGATLERRAKTLSPGKNGVYKDVAFGAVENPEYLPQGGAAQPPAPAFSP 1215  
QY 508 AFNLYVWODPPREGAPSPSTFKGTPTAENPEYLGIDVP 546  
Db 1216 AFNLYVWODPPREGAPSPSTFKGTPTAENPEYLGIDVP 1254

## RESULT 6

US-08-356-786-2  
Sequence 2, Application US/08356786  
Patent No. 5877305

## GENERAL INFORMATION:

APPLICANT: Huston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/356,786

## CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-356-786-2

Query Match 64.3%; Score 2405.5; DB 2; Length 1255;  
Best Local Similarity 41.1%; Pred. No. 4.8e-179;  
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;  
QY 35 LANGAASSTGVCTGDMKRLRPAPEHRLDMLRLYLQGCQVVGQNLLETLYPTNASLSFLQ 94

Db 16 LPFGAASSTGVCTGDMKRLRPAPEHRLDMLRLYLQGCQVVGQNLLETLYPTNASLSFLQ 75  
QY 95 DIOEVGVYLIHNOVROVLORLRIYRGTOLEEDNVALVLNGDPLNNTTPTGASPG 154  
Db 76 DIOEVGVYLIHNOVROVLORLRIYRGTOLEEDNVALVLNGDPLNNTTPTGASPG 135  
QY 155 GLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOALTLIDTNSRACHPC 214  
Db 136 GLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOALTLIDTNSRACHPC 135  
QY 215 SPMKSGRCWGESSESDCQSLTRFYVACAGCARCGPLPTDCHCQCAAGCTGPRHSDCLAC 274  
Db 196 SPMKSGRCWGESSESDCQSLTRFYVACAGCARCGPLPTDCHCQCAAGCTGPRHSDCLAC 255  
QY 275 LHFHNSGICELHCPALVTYTTDTFESMPNPEGRTFGASCVTACPNYLTSDVGS----- 329  
Db 256 LHFHNSGICELHCPALVTYTTDTFESMPNPEGRTFGASCVTACPNYLTSDVGSCTLVC 315  
QY 330 ----- 329  
Db 316 PLHNOEYTAEDGTORCEKSKPCARVCYGLGMELREAVTSAHQEFAGCKKIFGSLA 375  
QY 330 ----- 329  
Db 376 FLPESEFGDPSANTAPLOPOLQVFTLEBITGYLYISAMPDLPDLVSFONLQVIRGRI 435  
QY 330 ----- 329  
Db 436 LHNQAVSLTQGLGSLWGLSLRELGGSLALHNNHNLFCVHTVPMDFLRPHQALH 495  
QY 330 ----- 329  
Db 496 TANRDECEYEGELACHQLCARGHCWGPPTQCVNSQFLRGCEVECHVLQGLPREV 555  
QY 330 ----- 329  
Db 556 NARHCLPHECOPONGSVTCFEPEDQVCAHYKDPFCVARCPSGVKPDLSYPTWK 615  
QY 330 ----- 329  
Db 616 FPDEBGACQPCPINCTHSCVDLDDKGCAPQARASPLTISAVGILVYLVGVGILI 675  
QY 330 ----- 329  
Db 676 KRRQOKIRKYMRLLOETELVEPLTPSGAMPQAMRLKETELRKVKVLSGAFGVY 735  
QY 330 ----- 329  
Db 736 KGIMIPDGENVKIPAIKVLRENTSPRANKELIDEAVYVAGVSPYVSRILGICLSTVQ 795  
QY 330 ----- 329  
Db 796 LVTQLMPYGLLDHVHNRGRGLSQDLLNMCQIANGMSYLEDBVLVHRDLAARVLYKS 855  
QY 330 ----- 329  
Db 856 PNHVKITDFGLARLDIDETEHADGKVPKIMMALESILRRFTHQSDVMSYGVYVWEL 915  
QY 330 ----- 329  
Db 916 MTEGAKPYDGIPIAREIPDLLEKGERLPPICTIDYIMVYKMMIDSECRPFRELVS 975  
QY 330 ----- 329  
Db 976 FSHMARDQRFVVIQNEEDLGPASPLDSTFYRSLLEDDMDGLVDABEYLVPQGFECPPD 1035  
QY 330 --GAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMAA 387  
Db 1036 APAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMAA 1095  
QY 388 KGQSLPTHPSPLOKRSSEPTVPLPSEIDGYAPLTCSPQPEYVNOQDVPQPPSPREG 447



Db 1096 KGLSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPQPEVYNQDVPARPPSPREG 1155  
Qy 448 PLPAAPAGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQPHPPAPSP 507  
Db 1156 PLPAAPAGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQPHPPAPSP 1215  
Qy 508 AFDNLVYWDODPPERGAPSPSTFKGPTAENPEYLGLDVP 546  
Db 1216 AFDNLVYWDODPPERGAPSPSTFKGPTAENPEYLGLDVP 1254

RESULT 7  
US-08-466-6808-68  
; Sequence 68, Application US/084666808  
; Patent No. 6075122  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,6808  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 68:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-466-6808-68

Query Match 64.3%; Score 2405.5; DB 3; Length 1255;  
Best Local Similarity 41.1%; Pred. No. 4.8e-179;  
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

Qy 35 LARGAASOVCTGDMKRLPASPETHDMLRHLVGGCOVVOGNLELYLPTNASTLQ 94  
Db 16 LPPGAASOVCTGDMKRLPASPETHDMLRHLVGGCOVVOGNLELYLPTNASTLQ 75  
Qy 95 DIOEVQGYVLAHNOVROVPLQRLRIVRGTOLEFEDNVALAVLDNDPILNNTPTVYGASPG 154  
Db 76 DIOEVQGYVLAHNOVROVPLQRLRIVRGTOLEFEDNVALAVLDNDPILNNTPTVYGASPG 135  
Qy 155 GLREQLSLSTFEILKGVLIQRNPOLCYQDTILMKDIFHKNNQLALLTIDNRSACHPC 214  
Db 136 GLREQLSLSTFEILKGVLIQRNPOLCYQDTILMKDIFHKNNQLALLTIDNRSACHPC 195  
Qy 215 SPMCKSGKMGSSSEDCSLRTVYAGAGCARCKGLPTDCHEQCAAGCTGPKHSDCLAC 274  
Db 196 SPMCKSGKMGSSSEDCSLRTVYAGAGCARCKGLPTDCHEQCAAGCTGPKHSDCLAC 255  
Qy 275 LHFNHSGICEIHCALVYNTDTFESMPNPEGRTYGASCYATACPYNYLSTDVGS----- 329

Db 256 LHFNHSGICEIHCALVYNTDTFESMPNPEGRTYGASCYATACPYNYLSTDVGSCLTVLC 315  
Qy 330 ----- 329  
Db 316 PLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREYRAVTSANIOEFACCKKIFGSLA 375  
Qy 330 ----- 329  
Db 376 FLPESEFGDASPANTAPLOPEQLOVFETLEITGYLYISAMPDSLPLDSYFONQYIRGRI 435  
Qy 330 ----- 329  
Db 436 LHNGAYS/LTQGLISWLGRLSLRELGSGLAIHHNTHLCFVHTVPMQDLFRNPHQALLH 495  
Qy 330 ----- 329  
Db 496 TANRDECEVGEGLACHQLCARCHCWGPPOCVNCSQFLRGQCEVCRVLOGLPREYV 555  
Qy 330 ----- 329  
Db 556 NARHCLPCHPECOQONSVTGFCEPADQVACAHYKDPFCVAPRPSGVKRPDLSTYMPYWK 615  
Qy 330 ----- 329  
Db 616 FPDEGACQPCPINCTHSCVDLDKGCPEQRASPLTISAVVGLLVVGVFGILI 675  
Qy 330 ----- 329  
Db 676 KRROQKIRKYTMRLLOETELVEPLTPSGAMPNOAMRILKETELKVKVLSGSAFGTVY 735  
Qy 330 ----- 329  
Db 736 KGIWTPDENKIPVAVKVLRENTSPKANKELDEAVYMAVGSPYSRLIGICLTSTVQ 795  
Qy 330 ----- 329  
Db 796 LVTLMPYGLLDHVENRGRIGSODLLNMCQIAKMSYLEDVRLVHRDLARNVLVKS 855  
Qy 330 ----- 329  
Db 856 PNHYKITDFGLARLLDIDETEVHADGKVPKIMMALESILRRFTHOSDWSYGVYWEI 915  
Qy 330 ----- 329  
Db 916 MTFGAKPYDGIAPAREIPDLLEKGERLPQPICTIDVYIMVYKCMNIDSECRPRRELVE 975  
Qy 330 ----- 329  
Db 976 FSRMARDPQRFVLIQNEDLGASPPLDSTFYRSLEDDMGDLVDAEYLVPPQGFPCDP 1035  
Qy 330 --GAGGVNHHNRSSSTSGGGDLTGLSEPEEARPRPLAPBSAGSDVFDGDKMAA 387  
Db 1036 APGAGGVNHHNRSSSTSGGGDLTGLSEPEEARPRPLAPBSAGSDVFDGDKMAA 1095  
Qy 388 KGLSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPQPEVYNQDVPARPPSPREG 447  
Db 1096 KGLSLPTHDSPLQRYSEDPVLPSETDGYVAPLTCSPQPEVYNQDVPARPPSPREG 1155  
Qy 448 PLPAAPAGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQPHPPAPSP 507  
Db 1156 PLPAAPAGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQPHPPAPSP 1215  
Qy 508 AFDNLVYWDODPPERGAPSPSTFKGPTAENPEYLGLDVP 546  
Db 1216 AFDNLVYWDODPPERGAPSPSTFKGPTAENPEYLGLDVP 1254

RESULT 8  
US-08-484-438-8  
; Sequence 8, Application US/08484438  
; Patent No. 581098  
; Patent No. 581098 5780031

```

GENERAL INFORMATION:
APPLICANT: Plowman, Gregory D.
APPLICANT: Cuioussou, Jean-Michel
APPLICANT: Shoyab, Mohammed
APPLICANT: Slegall, Clay B.
APPLICANT: Hellster m, Ingegerd
APPLICANT: Hellster m, Karl E.
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennle & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-8

Query Match      63.8%; Score 2385.5; DB 2; Length 1255;
Best Local Similarity 41.0%; Pred. No. 1.8e-177;
Matches 509; Conservative 0; Mismatches 2; Indels 729; Gaps 3;

```

```

|||||
Db 256 LHFHNSGICELHCPALVTYNTDTFESMPNEGRTFGASCVTACPYNYLSTDVGSCTLYC 315
QY 330 ----- 329
Db 316 PLHNOEVTAEADGTORCEKCKPCARVCYGLGMEHLREVRVTSANTIOEFACKKIRGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGDPASNTAPLQRPQLOVFETLEETGYLYISAMPDSLPLDSVONQVINGRI 435
QY 330 ----- 329
Db 436 LHNQAVSLTLOGIGISWLGRLSRELGSGLALIHNTHLCFVHTVPWDQLFRNHQALLH 495
QY 330 ----- 329
Db 496 TANPEDECVBGLACHQCLARRALLSGPTQVCNCSQFLRGQECVECAVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCAHAHYKDPFCVAPCPGVAHPDLSYMPIWK 615
QY 330 ----- 329
Db 616 FPDEGACOPCPINCSTHSCVDLDDKGCAPABORASPLTSIVSAVGLLVVLGVFGILI 675
QY 330 ----- 329
Db 676 KRROQKIRKTMRLDQETLEVEPLTPSGAMPNOAQRILKETELRKVKYLSGAGCTYV 735
QY 330 ----- 329
Db 736 KGIIWPDGENVKIPVAKIVRENTSPKANKEIIDEAYVAGVSPYVSRIGLIGLSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGLLDHVRNKRRLQSOLLNKCQIANKMSYLEDRVLRDLAARNLVYS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLDIDETEVHADGKVPKMMALESILRRRTHQSDVMSVGYVMEL 915
QY 330 ----- 329
Db 916 MTEGAKPYDIPAREIPDLLEKGERLPQPICTIDVYIMVKCMMIDSECRPRELVS 975
QY 330 ----- 329
Db 976 FSRMARDPQRFVYIQNEIDLGPASPLDSTFYRSLLEDDMDGLDYDAEYLVPOGFGCPDP 1035
QY 330 -GAGGAVHHRHSSSTRSGGDLTLGLEPSEBAPRSPAPSEAGSDVFDGDLGAA 387
Db 1036 APGAGGVHHRHSSSTRSGGDLTLGLEPSEBAPRSPAPSEAGSDVFDGDLGAA 1095
QY 388 KGIDSLPTHPSPLOKRSSEPTVPPLPSETDGYVAPRLCSQPREVYNQDVRPQRPSPREG 447
Db 1096 KGIDSLPTHPSPLOKRSSEPTVPPLPSETDGYVAPRLCSQPREVYNQDVRPQRPSPREG 1155
QY 448 PLPAARAGATLBRAKTLSPGKNGVYVDFAFGAVENPEYLVPOGGAAPQHPPPAFSP 507
Db 1156 PLPAARAGATLBRAKTLSPGKNGVYVDFAFGAVENPEYLVPOGGAAPQHPPPAFSP 1215
QY 508 AFDNLVYWDQDPPERGAAPSTFKGTP -ADNPEYLGIDVP 546
Db 1216 AFDNLVYWDQDPPERGAAPSTFKGTPVAINPEY -GLDVP 1254

RESULT 9
US-09-146-283-4
: Sequence 4, Application us/09146283
: Patent No. 5976546
: GENERAL INFORMATION:

```

```

: APPLICANT: Laus, Relner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Compositions
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/146,283
: FILING DATE: 03-SEPT-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010.21
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0880
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 782 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
: US-09-146-283-4

Query Match      59.58; Score 2224; DB 2; Length 782;
Best Local Similarity 59.28; Pred. No. 3.6e-165;
Matches 478; Conservative 18; Mismatches 104; Indels 208; Gaps 20;

QY 35 LARGAATQVCTGDMKRLRLPASPETHLMDLRHLTYGCGVQVGNLELYLPTNASLSFLQ 94
D 16 LPPGASATQVCTGDMKRLRLPASPETHLMDLRHLTYGCGVQVGNLELYLPTNASLSFLQ 75
QY 95 DIOEVQGVYLLAHNQVQVPLQRLRLRYRGTLQFEDNVALAVLDNGDPLNNTPTVTGASPG 154
D 76 DIOEVQGVYLLAHNQVQVPLQRLRLRYRGTLQFEDNVALAVLDNGDPLNNTPTVTGASPG 135
QY 155 GLRELQLRLSLLEILKGVLLIQRNPQLCYQDTILMKDIFHKNNQALALTLIDNRSACHP 214
D 136 GLRELQLRLSLLEILKGVLLIQRNPQLCYQDTILMKDIFHKNNQALALTLIDNRSACHP 195
QY 215 SPMCKGRCMGESSEDCOSLTRTVCAGGACARCKGRLPTDCHEOCACACTGPKHSDIAC 274
D 196 SPMCKGRCMGESSEDCOSLTRTVCAGGACARCKGRLPTDCHEOCACACTGPKHSDIAC 255
QY 275 LHFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCYACPYNYLSTDVSGAGM 334
D 256 LHFHNSGICELHCPALVYNTDTFESMPNPEGRYTFGASCYACPYNYLSTDVSGAGM 313
QY 335 VHHHRSSSTSGGCD-----LTLGLEPSEEARPSPLAPSEAG----- 374
D 314 VCPHNOEVTVAEDGTQRCCKSCPCARVCYGLGMEHLREVAVYASANTQERAGCKIIFGS 373
QY 375 -----SDVFDCD----- 381
D 374 LAFLEPESDGPASNTAPLOPBOLOVFETLEITGYLYISAMPDLSPLSFONIQVTRG 433
QY 382 -----LGMGA- KGLQSL-----PTH-----DPSFLQRYSEDP----- 408

```

```

D 434 RILHNGAVSLTLQIGISLGLRLSLRELGLALIHNTHLCFVHTVWMDLFRNPHQAL 493
QY 409 --TYPLESET---DGYVAPLTC-----POP-----E 430
D 494 LHTANRPEDCEVGGELACHQICARGHCWPGPTQCVNCOSQPLRGOECVEECVLYGLPRE 553
QY 431 YVN-----OPVROPSPREGPLPARAPAGATLERAKTILSPKNGVAVDAFAGAV 483
D 554 YVNARHCLPCHPECOPO-----NG---SYCFGEA 581
QY 484 EN---PEYLPQGAAPQPPHPPAFSPAFDNLVYWDQDPPEGA-PPSTF----- 529
D 582 DQVACAAYHKDPPPCVA---RCPGKVPKDLSTYPIW-KFPDEGACQPCINCTASCVDL 637
QY 530 --KGTPTNENEYVGLDVPAAPARSPSPQPMHVAIOEARLNLSTDPTAAEMNET 587
D 638 DDKCPARQARSP---TSLAPARSPSPQPMHVAIOEARLNLSTDPTAAEMNET 694
QY 588 VEVISSEMDLQPTCLQRLRLRYKGLRSLTKLKGPLTMASHYKQHCPTPETSCATQ 647
D 695 VEVISSEMDLQPTCLQRLRLRYKGLRSLTKLKGPLTMASHYKQHCPTPETSCATQ 754
QY 648 ITFESFENLKDPLVLPFDQWEPVOE 675
D 755 ITFESFENLKDPLVLPFDQWEPVOE 782

```

```

RESULT 10
US-08-579-823A-4
: Sequence 4, Application US/08579823A
: Patent No. 6080409

```

## GENERAL INFORMATION:

```

: APPLICANT: Laus, Relner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Composition and Method
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto

```

```

: STATE: CA
: COUNTRY: USA

```

## COMPUTER READABLE FORM:

```

: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/579,823A
: FILING DATE: 03-DEC-1998

```

## CLASSIFICATION: 536

## ATTORNEY/AGENT INFORMATION:

```

: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010
: TELECOMMUNICATION INFORMATION:

```

```

: TELEFAX: 650-324-0960

```

```

: INFORMATION FOR SEQ ID NO: 4:

```

```

: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 782 amino acids

```

```

: TYPE: amino acid

```

```

: TOPOLOGY: linear

```

```

: MOLECULE TYPE: protein

```

```

: HYPOTHEICAL: NO

```

```

: ORIGINAL SOURCE:

```

```

: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
: US-08-579-823A-4

```

Query Match 59.5%; Score 2224; DB 3; Length 782;  
 Best Local Similarity 59.2%; Pred. No. 3.6e-165;  
 Matches 478; Conservative 18; Mismatches 104; Indels 208; Gaps 20;

```

OY 35 LARGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVAGNLELTYLPTNASSLFQ 94
DB 16 LPRGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVAGNLELTYLPTNASSLFQ 75
OY 95 DIOEVQGVYLIANNOVQVFLQRLRIYRGTOLEFEDNALAVLNDGDP LNNTTPTVGASPG 154
DB 76 DIOEVQGVYLIANNOVQVFLQRLRIYRGTOLEFEDNALAVLNDGDP LNNTTPTVGASPG 135
OY 155 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 214
DB 136 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 195
OY 215 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHECCAGCTGPKHSDDLAC 274
DB 196 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHECCAGCTGPKHSDDLAC 255
OY 275 LHFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNYLSTDVSGAGM 334
DB 256 LHFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNYLSTDVSGCT--L 313
OY 335 VHHRRSSSTRSGGD-----LTLGLEPSEEA PRSPLAPSGAG----- 374
DB 314 VCPHNOEVAEDGTORCEKSRPCARVCYGLGMEHLREVRAYTSANIOEFAGCKIIFS 373
OY 375 ----SDVFDD----- 381
DB 374 LAFLEPSFDDDPASNTAPLOPEQLQVETLEITGYLISAMPDSDLVSFONLQYIRG 433
OY 382 -----LGMGA--KGLOSL-----PTN---DPSPLQRYSEDP--- 408
DB 434 RILHNGAYSLTLOGIGLSMWGLSLRELGSGLALIHNTHLCEFNHYVPMQOLFNPQAL 493
OY 409 --TPVLPSEI--DGYVAPLTC-----PQ-----E 430
DB 494 LHTANREDECVGELGACHOLGACHGICMGPGPTQCVNCSQFLRGQECVECRVLQGLPRE 553
OY 431 YVN-----QPDVROPQPPRREGPLPAARPAATLERAKTLSPGKGVYKDVAFGAV 483
DB 554 YVANRHCLPCHPECPQ-----NG-----SYTCFPEA 581
OY 484 EN---PEVLTPOGGAAPORHPAPASPAFDNLYWDDPRERGA--PPSTF----- 529
DB 582 DDCVACAHYKDRPECVA---RCPSGVKRDLSYMPIM--KPFDEBAGACQPCPINCTHSCVDL 637
OY 530 --KGTPTAENPEYLGIDVPAARARSPSTQPRWEHNAIOEARLLNLSDTAEMNET 587
DB 638 DDGCPRAEQNASPL---TSLEAPARSPSPSTQPRWEHNAIOEARLLNLSDTAEMNET 694
OY 588 VEYISEAFDQEPCTCLQTRLELYKQGLRGSLTKLPGUTYMAASHYKOHCPPTSCATO 647
DB 695 VEYISEAFDQEPCTCLQTRLELYKQGLRGSLTKLPGUTYMAASHYKOHCPPTSCATO 754
OY 648 ITTFESFENKIDFLVLIIPDCWEPYOE 675
DB 755 ITTFESFENKIDFLVLIIPDCWEPYOE 782

```

RESULT 11  
 US-09-344-195-4  
 ; Sequence 4, Application us/09344195  
 ; Patent No. 6210662  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Laus, Reiner  
 ; Ruegg, Curtis L.  
 ; Mu, Hongyu  
 ; TITLE OF INVENTION: Immunostimulatory Compositions  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave. Suite 250  
 City: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94306

COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/344,195  
 FILING DATE: 24-Jun-1999  
 CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/146,283  
 FILING DATE: 03-SEPT-1998

ATTORNEY/AGENT INFORMATION:  
 NAME: Judge, Linda R.  
 REGISTRATION NUMBER: 42,702  
 REFERENCE/DOCKET NUMBER: 7636-0010.21  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-324-0880  
 TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 782 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: homo sapiens  
 INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein: Fig. 8  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-344-195-4

Query Match 59.5%; Score 2224; DB 4; Length 782;  
 Best Local Similarity 59.2%; Pred. No. 3.6e-165;  
 Matches 478; Conservative 18; Mismatches 104; Indels 208; Gaps 20;

```

OY 35 LARGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVAGNLELTYLPTNASSLFQ 94
DB 16 LPRGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVAGNLELTYLPTNASSLFQ 75
OY 95 DIOEVQGVYLIANNOVQVFLQRLRIYRGTOLEFEDNALAVLNDGDP LNNTTPTVGASPG 154
DB 76 DIOEVQGVYLIANNOVQVFLQRLRIYRGTOLEFEDNALAVLNDGDP LNNTTPTVGASPG 135
OY 155 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 214
DB 136 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 195
OY 215 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHECCAGCTGPKHSDDLAC 274
DB 196 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHECCAGCTGPKHSDDLAC 255
OY 275 LHFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNYLSTDVSGAGM 334
DB 256 LHFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNYLSTDVSGCT--L 313
OY 335 VHHRRSSSTRSGGD-----LTLGLEPSEEA PRSPLAPSGAG----- 374
DB 314 VCPHNOEVAEDGTORCEKSRPCARVCYGLGMEHLREVRAYTSANIOEFAGCKIIFS 373
OY 375 ----SDVFDD----- 381
DB 374 LAFLEPSFDDDPASNTAPLOPEQLQVETLEITGYLISAMPDSDLVSFONLQYIRG 433
OY 382 -----LGMGA--KGLOSL-----PTN---DPSPLQRYSEDP--- 408
DB 434 RILHNGAYSLTLOGIGLSMWGLSLRELGSGLALIHNTHLCEFNHYVPMQOLFNPQAL 493

```

QY 409 --TVLPSET---DGYVAPLTC-----POP-----E 430  
DB 494 LHTARPEDECGBGLACHQCARHCWGPGPTQCVCNCSQFLRGQECVEECVLOGLPRE 553  
QY 431 YVN-----OPDVRPPSPREGPLPARPAGATLERAKTSLPSKNGVYKDVAFGGAV 483  
DB 554 YVNARHCLPCHECCPQ-----NG-----SVTCFGEPA 581  
QY 484 EN-----PEYLRPGGAAPQPHPPAPAFDNLVYWDODPPERGA-PPSTF----- 529  
DB 582 DQVACAHAKDPPEFVA---KCPSSVKPDLSTYPIW-KPDEGACQCPPLNCTHSCVDL 637  
QY 530 --KGTPTAENPEYGLDVPAPAPASPSPTQPMENAVNIOEARRLNLSRDTAEMNET 587  
DB 638 DDKGCPAGORASPL---TSLEAPARSPSPQPMENAVNIOEARRLNLSRDTAEMNET 694  
QY 588 VEVISMFDLQEPCTQLRLLEYKQGLRGLSLTKLKGPLTMASHYKQHCPTPTSCATQ 647  
DB 695 VEVISMFDLQEPCTQLRLLEYKQGLRGLSLTKLKGPLTMASHYKQHCPTPTSCATQ 754  
QY 648 ITTFSEFKNLKDPLVYIPDCMPVQE 675  
DB 755 ITTFSEFKNLKDPLVYIPDCMPVQE 782

## RESULT 12

US-09-630-155-2  
; Sequence 2, Application US/09630155  
; Patent No. 6414130  
; GENERAL INFORMATION:  
; APPLICANT: Doherty, Joni Kristin and Gail M. Clinton  
; TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DAVIS WRIGHT TREMAINE LLP  
; STREET: 1501 Fourth Avenue, 2600 Century Square  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: PC compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/630,155  
; FILING DATE: 16-Jan-2001  
; CLASSIFICATION: <unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davison, Barry L.  
; REGISTRATION NUMBER: 47,309  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206 628-7621  
; TELEFAX: 206 628-7699  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 419  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: polypeptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-630-155-2

Query Match 43.0%; Score 1607; DB 4; Length 419;  
Best Local Similarity 77.2%; Pred. No. 2e-117;  
Matches 318; Conservative 9; Mismatches 55; Indels 30; Gaps 7;  
QY 35 LARGAASVQVCTGDMKRLRLPASPEHLDMLHLVQGGCVVQGNLELYLPNASTSLQ 94  
DB 16 LPPGAASVQVCTGDMKRLRLPASPEHLDMLHLVQGGCVVQGNLELYLPNASTSLQ 75

QY 95 DIOEVGYVILAHNOVROVPIORLIRVGTOLFEDNVALLAVLNDGDPPLNNTTPVTGASPG 154  
DB 76 DIOEVGYVILAHNOVROVPIORLIRVGTOLFEDNVALLAVLNDGDPPLNNTTPVTGASPG 135  
QY 155 GLRELQLRSLTEILKGVLIORNPOLCYODTILMDIFRKNQALTLTIDTRSRACHPC 214  
DB 136 GLRELQLRSLTEILKGVLIORNPOLCYODTILMDIFRKNQALTLTIDTRSRACHPC 195  
QY 215 SPMCKGSRCKGSSSDCCSLRTVCAGGACRKGPLPTDCCHEQCAAGCTGKRHSDCLAC 274  
DB 196 SPMCKGSRCKGSSSDCCSLRTVCAGGACRKGPLPTDCCHEQCAAGCTGKRHSDCLAC 255  
QY 275 LHFNSGICELCPALVYNTDTPESMPNPEGRTYFGASCYACPYNTLSTVSGAGAG 334  
DB 256 LHFNSGICELCPALVYNTDTPESMPNPEGRTYFGASCYACPYNTLSTVSGAGAG 313  
QY 335 VHHRRSSSTRSGGDLTLGLEPSEEARPS--PLAPSGAGSDVFDGLGMAAKGLQS 392  
DB 314 VCPRLNQEVTADG---TQRCCKSKPCARGTHSLPRPA-----VPVLRMQP 360  
QY 393 LPTHPSPLOKRSQD-----PTVPLPSETDGYVAPLTCSPQPEYVQNP 436  
DB 361 GPAPHVLSFLRPSWDLVSAFYSLPLAPL-SPTSPVSPISVSGRGPD---PD 497

## RESULT 13

US-08-422-108-1  
; Sequence 1, Application US/08422108  
; Patent No. 6015567  
; GENERAL INFORMATION:  
; APPLICANT: Hudziak, Robert M.  
; APPLICANT: Shepard, H. Michael  
; APPLICANT: Ullrich, Axel  
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Winpatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/422,108  
; FILING DATE: 14-Apr-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/355460  
; FILING DATE: 13-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/048346  
; FILING DATE: 15-APR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/354319  
; FILING DATE: 19-MAY-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M  
; REGISTRATION NUMBER: 00,000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 624 amino acids  
; TYPE: Amino Acid

TOPOLOGY: Linear  
US-08-422-108-1  
Query Match 42.7%; Score 1597; DB 3; Length 624;  
Best Local Similarity 79.8%; Pred. No. 2,1e-116;  
Matches 308; Conservative 6; Mismatches 34; Indels 38; Gaps 4;

QY 41 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 100  
DB 1 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 60  
QY 101 GYVLIANOVROVPLOLRIRYRGTOLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 160  
DB 61 GYVLIANOVROVPLOLRIRYRGTOLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120  
QY 161 LRSLEILKGGVLIQORNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPMCKG 220  
DB 121 LRSLEILKGGVLIQORNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPMCKG 180  
QY 221 SRCGSESEDCQSLTRFYVAGGACGKGPLPTDCHEQCAAGCTGPRHSDCLACLFHNHS 280  
DB 181 SRCGSESEDCQSLTRFYVAGGACGKGPLPTDCHEQCAAGCTGPRHSDCLACLFHNHS 240  
QY 281 GICELHCPALVTYNTDFEESMPNEGRTYGASCTACPYNYLSTDVSGAGGVHHRHR 340  
DB 241 GICELHCPALVTYNTDFEESMPNEGRTYGASCTACPYNYLSTDVSGCT--LVCPHLN 298  
QY 341 SSSSTRSGCD-----LTGLPESEEBAPRSPPLAPSEAG-----SD 376  
DB 299 QEVTAEDGTQCEKCKSPCARVCGTGMENHREAVTASNIOGFACKKIFGSLAFLE 358  
QY 377 VFDGDLGMAKGLQSLPTHDPSFLQ 402  
DB 359 SFDGD-----PASNTAPLQ 372

RESULT 14  
US-08-422-734-1  
Sequence 1, Application US/08422734  
Patent No. 6333169  
GENERAL INFORMATION:  
APPLICANT: Hudziak, Robert M.  
APPLICANT: Ullrich, Axel  
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/422,734  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/422108  
FILING DATE: 14-Apr-1995  
APPLICATION NUMBER: 08/355460  
FILING DATE: 13-Dec-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/048346  
FILING DATE: 15-Apr-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/354319  
FILING DATE: 19-May-1989

ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 554C2D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-422-734-1  
Query Match 42.7%; Score 1597; DB 4; Length 624;  
Best Local Similarity 79.8%; Pred. No. 2,1e-116;  
Matches 308; Conservative 6; Mismatches 34; Indels 38; Gaps 4;

QY 41 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 100  
DB 1 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 60  
QY 101 GYVLIANOVROVPLOLRIRYRGTOLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 160  
DB 61 GYVLIANOVROVPLOLRIRYRGTOLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120  
QY 161 LRSLEILKGGVLIQORNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPMCKG 220  
DB 121 LRSLEILKGGVLIQORNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPMCKG 180  
QY 221 SRCGSESEDCQSLTRFYVAGGACGKGPLPTDCHEQCAAGCTGPRHSDCLACLFHNHS 280  
DB 181 SRCGSESEDCQSLTRFYVAGGACGKGPLPTDCHEQCAAGCTGPRHSDCLACLFHNHS 240  
QY 281 GICELHCPALVTYNTDFEESMPNEGRTYGASCTACPYNYLSTDVSGAGGVHHRHR 340  
DB 241 GICELHCPALVTYNTDFEESMPNEGRTYGASCTACPYNYLSTDVSGCT--LVCPHLN 298  
QY 341 SSSSTRSGCD-----LTGLPESEEBAPRSPPLAPSEAG-----SD 376  
DB 299 QEVTAEDGTQCEKCKSPCARVCGTGMENHREAVTASNIOGFACKKIFGSLAFLE 358  
QY 377 VFDGDLGMAKGLQSLPTHDPSFLQ 402  
DB 359 SFDGD-----PASNTAPLQ 372

RESULT 15  
US-08-414-417B-69  
Sequence 69, Application US/08414417B  
Patent No. 5801005  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disis, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/414,417B

FILING DATE: 31-MAR-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
REFERENCE/DOCKET NUMBER: 920010.448C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEO ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 580 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-414-417B-69

Query Match 31.6%; Score 1183; DB 1; Length 580;  
Best Local Similarity 95.6%; Pred. No. 3.8e-84;  
Matches 219; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

QY 318 CPYVLTSDVSGAGWVHHRRSSSTRSGGDLTLGLEPSEERAPRSPPLAPSGAGSDV 377  
DB 357 CP-----DPAAGAGWVHHRRSSSTRSGGDLTLGLEPSEERAPRSPPLAPSGAGSDV 410  
QY 378 FDGDLGMAKGLQSLPTHDSPLOKRSSEDPVPLPSETDGYVAPLTCSPOPEYVNOPDV 437  
DB 411 FDGDLGMAKGLQSLPTHDSPLOKRSSEDPVPLPSETDGYVAPLTCSPOPEYVNOPDV 470  
QY 438 RPQPPSPREGPLPAARPAATLERAKTLSPGKNGVAKDVFAFGAVENPEYLTPOGGAAP 497  
DB 471 RPQPPSPREGPLPAARPAATLERAKTLSPGKNGVAKDVFAFGAVENPEYLTPOGGAAP 530  
QY 498 QPHPPAPSPAFDNLVYWDODPERGAPSPSTFKGPTAENPEYLGIDVP 546  
DB 531 QPHPPAPSPAFDNLVYWDODPERGAPSPSTFKGPTAENPEYLGIDVP 579

Search completed: April 28, 2003, 13:43:11  
Job time : 34.1923.secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:49 : Search time 46.1538 Seconds  
(without alignments)  
3080.404 Million cell updates/sec

Title: US-09-821-883-2  
Perfect score: 3739  
Sequence: 1 MRAAPLLARAAASLSIGFLF.....EPVQEGAPPPRAAHNNHHH 690

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_todent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rviro:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2124.5	56.8	1259	6 O18735	O18735 canis famli
2	1609.5	43.0	419	4 O9UK79	O9UK79 homo sapien
3	969	25.9	367	11 O8R2X1	O8R2X1 mus musculu
4	906	24.2	165	4 O14256	O14256 homo sapien
5	900.5	24.1	412	4 O8WYV0	O8WYV0 homo sapien
6	754.5	20.2	527	13 O90836	O90836 gallus gall
7	723	19.3	149	6 O9BG66	O9BG66 oryctolagus
8	717.5	19.2	643	11 O9ERV6	O9ERV6 mus musculu
9	717.5	19.2	655	11 O9WVF5	O9WVF5 mus musculu
10	717.5	19.2	1210	11 O9EP98	O9EP98 mus musculu
11	714.5	19.1	1209	11 O9QX70	O9QX70 rattus norv
12	704.5	18.8	478	11 O9ESES0	O9ESES0 rattus norv
13	697	18.6	331	4 O9BUD7	O9BUD7 homo sapien
14	664.5	17.8	1165	13 O9YH40	O9YH40 xiphophorus
15	654	17.5	144	6 O9GL44	O9GL44 macaca mlla
16	641.5	17.2	599	13 O9PSH2	O9PSH2 gallus gall

17	608	16.3	1328	13 P79754	P79754 fugu rubrip
18	545	14.6	144	6 O9MYK4	O9MYK4 ovis aries
19	543	14.5	1433	5 O9B1H9	O9B1H9 anophelis g
20	528	14.1	146	6 O8WN17	O8WN17 equus caball
21	515	13.8	152	6 O9SL10	O9SL10 equus caball
22	427	11.4	141	11 O8VH40	O8VH40 sigmodon thl
23	407	10.9	1137	13 O9W6F6	O9W6F6 gallus gall
24	398	10.6	150	6 O9BG64	O9BG64 oryctolagus
25	395	10.6	151	6 O9BG65	O9BG65 oryctolagus
26	377	10.1	1368	5 O23821	O23821 caenorhabdi
27	362	9.7	138	11 O99J91	O99J91 marmota mon
28	341	9.1	366	5 O26569	O26569 schistosoma
29	341	9.1	1717	5 O26566	O26566 schistosoma
30	331	8.9	334	5 O26567	O26567 schistosoma
31	331	8.9	342	5 O26568	O26568 schistosoma
32	323.5	8.7	1193	5 O9Y1X8	O9Y1X8 ephydratia f
33	291	7.8	1472	5 O9U5A8	O9U5A8 bombyx mori
34	289	7.7	1671	5 O9NJV5	O9NJV5 bioplmalari
35	274	7.3	1749	5 O8T0W6	O8T0W6 echinococcu
36	273.5	7.3	89	11 O88459	O88459 mus musculu
37	273.5	7.3	1418	13 O8UW83	O8UW83 parathichth
38	272.5	7.3	1418	13 O93457	O93457 scophthalmu
39	270.5	7.2	2144	5 O9VD94	O9VD94 drosophila
40	263	7.0	1358	13 O73798	O73798 xenopus lae
41	259.5	6.9	1412	13 O8UW84	O8UW84 parathichth
42	256.5	6.9	1362	13 O9PV24	O9PV24 xenopus lae
43	256	6.8	1369	13 O8UW85	O8UW85 parathichth
44	247	6.6	1368	13 O8UW85	O8UW85 parathichth
45	238	6.4	469	11 O63721	O63721 rattus norv

## ALIGNMENTS

### RESULT 1

O18735 PRELIMINARY: PRT: 1259 AA.  
ID O18735  
AC O18735;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Erdb-2.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_Taxid=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yokota H.;  
RT "CDNA cloning of erdb-2 from canine mammary gland."  
RL Submitted (OGF-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB008451; BAA23127.1;  
DR HSBP; P11362; IFC.  
DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR InterPro: IPR004019; YLP\_motif.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF00069; pkinase; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
DR Pfam: PF02757; YLP; 2.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART; SM00261; FU; 3.  
DR SMART; SM00219; TYRC; 1.  
DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW ATP-binding; Transferase; Tyrosine-protein kinase.  
SQ SEQUENCE 1259 AA; 137989 MW; E37364D49C4ACD46 CRC64;

Query Match 56.8%; Score 2124.5; DB 6; Length 1259;  
 Best Local Similarity 36.7%; Pred. No. 1e-149;  
 Matches 460; Conservative 18; Mismatches 44; Indels 731; Gaps 2;

```

QY 25 MLDRSVLAKELARGAATGCTGDMKRLPASPETHLDMRLHYGCCVVGQLELTLYL 84
DB 6 WCRMGILLALPSSAAGTCTGDMKRLPASPETHLDMRLHYGCCVVGQLELTLYL 65
QY 85 PTNASTFLDIOEVQGVLIANQVQVPLQRLRIYRGTQLEFEDNALVALNDGDLPLN 144
DB 66 PANASTSFLDIOEVQGVLIANSQVQVPLQRLRIYRGTQLEFEDNALVALNDGDLPLN 125
QY 145 TTPPTGASPGGLRELRLSLTEILKGVLIQRPOLCYOTIIMKDIFFHKNNOALTLID 204
DB 126 GIAPAPGAAGGLRELRLSLTEILKGVLIQRPOLCYOTIIMKDIFFHKNNOALTLID 185
QY 205 TNRBRACHPCSPMGKSGRGESEDCQSLTRTYCAGGACARCKPPLTDCHECCAGCT 264
DB 186 TNRFSACPPSPACKDACHGASSGDCQSLTRTYCAGGACARCKPPLTDCHECCAGCT 245
QY 265 GPKHSDCLACLFHNHSGICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLS 324
DB 246 GPKHSDCLACLFHNHSGICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLS 305
QY 325 TDVCS----- 329
DB 306 TDVGSCTLVCPLNQEVTAEDGTQRCCKSPCARVCYGLMEHLREVRVTSANIOEFA 365
QY 330 ----- 329
DB 366 GCKKIFSLAFIPESFPGDASNTAPLPQPOLRVFEALBEITGLYISANPDSLPLNSVF 425
QY 330 ----- 329
DB 426 QNLRVIRGLVLDGAVSLTLQGLISWGLRSLRELSSGLALIHNRNARLCFVHTVMDQL 485
QY 330 ----- 329
DB 486 FRNPQHALLHSANRPEECVGEGLACYPCAHGHCWGPPTQVCNCSQFLRGQCEVECRV 545
QY 330 ----- 329
DB 546 LQGLPREYVADRYCLPCHSECPQNGSVTCFSGADQCVACANHKDPFCVARGCPSGVKP 605
QY 330 ----- 329
DB 606 DLSEPMIKFADEGTCQPCPINCTHSCADLDEKGCRAQBPASVTSIIAAVVGILLAV 665
QY 330 ----- 329
DB 666 VGLVILILIRRKOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETELRKVKVL 725
QY 330 ----- 329
DB 726 GSGAFGVYKGIWIPDENKIPVAVKVLAKENTSPKANKELDEAVYVAGVSGSYVSRLL 785
QY 330 ----- 329
DB 786 GICLSTVOLVTOIMPYGCLLDHVHRGRGLSQDLLNMCVQAKGMYLEDVRLVHRL 845
QY 330 ----- 329
DB 846 AARNVLKSPNHVKITDGLARLLIDETEHADGKVPKIMMALESIPPRRFTHQSDVW 905
QY 330 ----- 329
DB 906 SYGVTVWELMTFGAKRPIDGIPAREIPDLLEKGERLPQPICTIDVYIMVCKMIDSECR 965
QY 330 ----- 329
DB 966 PRFRELVAEFSRMAPDRQRFVYIQNEDLGASPLDSTFYRSILEDDEDDMDGLVDAEVLVP 1025
QY 330 -----GAGGVVHHRHRSSTRSGGDLTLGLEPSEEBAPRSLAPSEGAGSDVFP 378
  
```

Query Match 43.0%; Score 1609.5; DB 4; Length 419;  
 Best Local Similarity 74.3%; Pred. No. 6.2e-112;  
 Matches 318; Conservative 7; Mismatches 42; Indels 61; Gaps 7;

```

DB 1026 QCGFCEPPTPGAGCTAHRHRSSTRNGGELTLGLEPSEEBKPSPLAPSEGAGSDVFP 1085
QY 379 DGDGMAAGAGLQSLPHNDSPLQRYSEDDPTVPLPSETDGVAVPLTCSPOPEYVNOPDR 438
DB 1086 DGDGMAAGAGLQSLPHNDSPLQRYSEDDPTVPLPSETDGVAVPLTCSPOPEYVNOPDR 1145
QY 439 POPSPREGLPAPAPAGATLER-----AKTLPKNGVYKDVAFAGAVENPEYLPPOG 493
DB 1146 POPPLEALEGLPAPAPAGATLERKPTLSPTLSFGKNGVYKDVAFAGAVENPEYLPARG 1205
QY 494 GAAPQHPHPAPSPAFONLYTMDODPPERCAPSTFEGTTPAENPEYLGLDVP 546
DB 1206 RAAQHPHPAPSPAFONLYTMDODPPERCAPSTFEGTTPAENPEYLGLDVP 1258

RESULT 2
Q9UK79 PRELIMINARY; PRT; 419 AA.
ID Q9UK79
AC Q9UK79;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Herstatin.
GN HER-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:99415951; PubMed:10485918;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted
RT autoinhibitor."
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Hennen W.D.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF177761; AAD56009.2;
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like.
DR Pfam: PF01030; Recep_L_domain; 1.
DR SMART: SM00261; Fu; 1.
SQ SEQUENCE 419 AA; 45472 MW; FECLBE347E2D030C CRC64;

Query Match 43.0%; Score 1609.5; DB 4; Length 419;  

  Best Local Similarity 74.3%; Pred. No. 6.2e-112;  

  Matches 318; Conservative 7; Mismatches 42; Indels 61; Gaps 7;


```

QY 35 LARGAATGCTGDMKRLPASPETHLDMRLHYGCCVVGQLELTLYLPTNASTSLQ 94
DB 16 LPRGAATGCTGDMKRLPASPETHLDMRLHYGCCVVGQLELTLYLPTNASTSLQ 75
QY 95 DIOEVQGVLIANQVQVPLQRLRIYRGTQLEFEDNALVALNDGDLPLNTPVTGASPG 154
DB 76 DIOEVQGVLIANQVQVPLQRLRIYRGTQLEFEDNALVALNDGDLPLNTPVTGASPG 135
QY 155 GLRELRLSLTEILKGVLIQRPOLCYOTIIMKDIFFHKNNOALTLIDTNRBRACHPC 214
DB 136 GLRELRLSLTEILKGVLIQRPOLCYOTIIMKDIFFHKNNOALTLIDTNRBRACHPC 195
QY 215 SPMCKSGRGESEDCQSLTRTYCAGGACARCKPPLTDCHECCAGCTGPKHSDCLAC 274
DB 196 SPMCKSGRGESEDCQSLTRTYCAGGACARCKPPLTDCHECCAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLS TDVSGAGM 334
DB 256 LHFHNSGICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLS TDVSGSC -L 313
QY 335 VHHHRHRSSTRSGGDLTLGLEPSEEBAPRSLAPSEGAGSDVFDGDLGMAAGLQSLP 394
  
```


```

```

Db 314 VCPHNOEVTAEDEG---TORCEKSKP-----CAROTHSLP 346
OY 395 THDPSPLQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVPRQ-----P 441
Db 347 ---PRP-----AAVPV-----PLRMQPGPAHVLFLRPSMDLVSAFYSLPLAP 387
OY 442 PSPREGPL 449
Db 388 LSPTSVP1 395

RESULT 3
OBR2X1 PRELIMINARY: PRT: 367 AA.
AC OBR2X1:
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 40.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
KW Hypothetical protein.
SQ SEQUENCE 367 AA; 40163 MW; 0BE03395F9E101B0 CRC64;

Query Match 25.9%; Score 969; DB 11; Length 367;
Best Local Similarity 80.0%; Pred. No. 3,1e-64;
Matches 180; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

OY 322 YLSTDVSGAGGVNHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASEGASGVDFDGD 381
Db 142 FSPDPALGTGSTAHRHRSSSARSNGGELTLGLEPSEEPSPPLASEGASGVDFDGD 201
OY 382 LGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVPRQ 441
Db 202 LAVGVTLGLOSLSPHDSPLQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVPRQ 261
OY 442 PSPREGPLPARRPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYTLAPAGTASOPHP 501
Db 262 PLTEGPPPPRPPAGATLERKTLPSPKNGVYKDVFAFGAVENPEYTLAPAGTASOPHP 321
OY 502 PPAESPAPFDNLVYWDQDPPERGAPSPFTFGTPTAENPEYTLGLDVP 546
Db 322 SPASPAPFDNLVYWDQNSSEGGPPSPFTFGTPTAENPEYTLGLDVP 366

RESULT 4
O14256 PRELIMINARY: PRT: 165 AA.
ID O14256:
AC O14256:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE C-erb B2/neu protein (Fragment).
GN C-ERB B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86070181; Pubmed=2999974;
RA Consens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,
RA Seeburg P.H., Libermann T.A., Schlessinger J., Francke U.,
RA Levinson A., Ullrich A.;
RA "Tyrosine kinase receptor with extensive homology to EGF receptor

```

```

RT shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=94000386; Pubmed=8104414;
RA Sakar F.H., Bail D.E., Li Y.W., Crisman J.D.;
RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)
RT gene.";
RL DNA Cell Biol. 12:611-615(1993).
DR EMBL; M95667; AAC37531.1; -.
FT NON_TER
SQ SEQUENCE 165 AA; 17327 MW; AOC113BA308BF46B CRC64;

Query Match 24.2%; Score 906; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.4e-60;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 373 AGSDVFDGDLGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYV 432
Db 1 AGSDVFDGDLGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYV 60
OY 433 NOPDVPRQPPSPRGGPLPARRPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYTLTPQ 492
Db 61 NOPDVPRQPPSPRGGPLPARRPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYTLTPQ 120
OY 493 GGAAPQHPAPPAPFDNLVYWDQDPPERGAPSPFTFGTPTAEN 537
Db 121 GGAAPQHPAPPAPFDNLVYWDQDPPERGAPSPFTFGTPTAEN 165

RESULT 5
OBRVVO PRELIMINARY: PRT: 412 AA.
ID OBRVVO:
AC OBRVVO:
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 44.7 kDa protein.
GN PP3659.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.Q., Qin W.X., Zhao X.T.,
RA Wan D.F., Gu J.R.;
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF318349; AL55856.1; -.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR001245; Tyr_Pkinase.
DR InterPro; IPR004019; YLP_motif.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF02757; YLP; 2.
DR Prodom; PD000001; Euk_Pkinase; 1.
DR SMART; SK00219; TykC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein.
SQ SEQUENCE 412 AA; 44702 MW; 034397FF3F27D2BC CRC64;

Query Match 24.1%; Score 900.5; DB 4; Length 412;
Best Local Similarity 75.3%; Pred. No. 4.7e-59;
Matches 183; Conservative 6; Mismatches 30; Indels 24; Gaps 5;

OY 322 YLSTDVSGAGGVNHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASEGASGVDFDGD 381
Db 142 FSPDPALGTGSTAHRHRSSSSTRSGGDLTLGLEPSEEAAPRSPPLASEGASGVDFDGD 201
OY 382 LGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVPRQ 441

```



RA Maible N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egr transcripts encoding truncated receptor  
 RT isoforms";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF275366; AAG28046.1;  
 DR EMBL: AF275364; AAG28046.1; JOINED.  
 DR EMBL: AF275365; AAG28046.1; JOINED.  
 DR MGI: 95294; Egr.  
 DR InterPro: IPR000345; CytC\_heme\_bind.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR002174; Furin-like.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L-like; 2.  
 DR SMART: SM00261; FU; 4.  
 DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 DR Receptor.  
 KW SEQUENCE 643 AA; 71476 MW; DEF220020C84911BI CRC64;  
 Query Match 19.2%; Score 717.5; DB 11; Length 643;  
 Best Local Similarity 46.8%; Pred. No. 3.8e-45;  
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;  
 QY 36 ARGAA--STOVCTGDMKRLPASPETHLDMRLHYQCGQVVGNGLELTPTNASTSFL 93  
 DB 20 AAGGALKEKKYCGQSTNRLTOLGTFEDHFLSLQRMYNCEVGLNLETTYQARNYDLSFL 79  
 QY 94 ODIOEGVGVLIANNOVQVPLQRLRIYRGQLEFEDNALVNDGDPDLNNTPTGASP 153  
 DB 80 KTIQEVAGYVLIALTVPRIPLNQLITIGNALVNTALALISN-----YGTNR 129  
 QY 154 GGLRELQRLSTLEILKGVLIQRIORPOLCYOPTIILMKDI-----FHNNOALATLIDTNSR 209  
 DB 130 TGLRELPRNRNLQELILIGAVRSNNPILCNMTIQRDIVQNVFMSMDL-----QSPHS 185  
 QY 210 ACHPSPCKSKRCWGESSEDCQSLTRTVACAGCA-RCKGPLPTDCHEGCAAGCTGPKH 268  
 DB 186 SCPKDPFPCPMGSCWGGEENCQKLTITCAQCSHRCGRSPSCCHNQCACGCTGPRE 245  
 QY 269 SDCLACLFHNSGICELCPALVYNTPTFESMPREGRYFGASCYACPNYISTDVG 328  
 DB 246 SDCLVQCFQDEATCKDTPPLMLNPTTYQMDVNPBESKYSFGATCKPCPNYVYTDHG 305  
 QY 329 S 329  
 DB 306 S 306  
 RESULT 9  
 Q9WVF5 PRELIMINARY; PRT; 655 AA.  
 AC Q9WVF5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Epidermal growth factor receptor (Epidermal growth factor receptor  
 DE isoform 3).  
 GN EGFR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LIVER.  
 RA Røtter J.L., Thredgill D.W., Danielsen A.J., Scheil C.,  
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
 RA Maible N.J.;  
 RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode  
 RT Carboxy-Terminal Truncated Receptors";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.

RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEVTAJ;  
 RA Røtter J.L., Thredgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
 RA Sinclair C.S., Pearlsall R.S., Green P.J., Yee D., Lampland A.L.,  
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,  
 RA Maible N.J.;  
 RT "Comparative genomic sequence analysis and isolation of human and  
 RT mouse alternative Egr transcripts encoding truncated receptor  
 RT isoforms";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=LIVER;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schiraldi L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barish G.,  
 RA Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kanlaya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombereis P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Welt C., Whittaker C., Winkler L.,  
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kotsuki S.,  
 RA Hayashizaki Y.;  
 RL "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AF124513; AAD44149.1;  
 DR EMBL: AF275366; AAG28047.1;  
 DR EMBL: AF275364; AAG28047.1;  
 DR EMBL: AF275365; AAG28047.1; JOINED.  
 DR EMBL: AK004944; BAB23688.1;  
 DR EMBL: AK004883; BAB23641.1;  
 DR EMBL: AK004911; BAB23662.1;  
 DR MGI: 95294; Egr.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR002174; Furin-like.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L-like; 2.  
 DR SMART: SM00261; FU; 3.  
 DR Receptor.  
 KW SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;  
 Query Match 19.2%; Score 717.5; DB 11; Length 655;  
 Best Local Similarity 46.8%; Pred. No. 3.9e-45;  
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;  
 QY 36 ARGAA--STOVCTGDMKRLPASPETHLDMRLHYQCGQVVGNGLELTPTNASTSFL 93  
 DB 20 AAGGALKEKKYCGQSTNRLTOLGTFEDHFLSLQRMYNCEVGLNLETTYQARNYDLSFL 79  
 QY 94 ODIOEGVGVLIANNOVQVPLQRLRIYRGQLEFEDNALVNDGDPDLNNTPTGASP 153  
 DB 80 KTIQEVAGYVLIALTVPRIPLNQLITIGNALVNTALALISN-----YGTNR 129  
 QY 154 GGLRELQRLSTLEILKGVLIQRIORPOLCYOPTIILMKDI-----FHNNOALATLIDTNSR 209  
 DB 130 TGLRELPRNRNLQELILIGAVRSNNPILCNMTIQRDIVQNVFMSMDL-----QSPHS 185  
 QY 210 ACHPSPCKSKRCWGESSEDCQSLTRTVACAGCA-RCKGPLPTDCHEGCAAGCTGPKH 268  
 DB 186 SCPKDPFPCPMGSCWGGEENCQKLTITCAQCSHRCGRSPSCCHNQCACGCTGPRE 245  
 QY 269 SDCLACLFHNSGICELCPALVYNTPTFESMPREGRYFGASCYACPNYISTDVG 328  
 DB 246 SDCLVQCFQDEATCKDTPPLMLNPTTYQMDVNPBESKYSFGATCKPCPNYVYTDHG 305



Query Match	19.1%;	Score 714.5;	DB 11;	Length 1209;
Best Local Similarity	46.7%;	Pred. NO. 1.5e-44;		
Matches 142;	Conservative 42;	Mismatches 105;	Indels 15;	Gaps 5

[illegible]

RESULT 12	
Q9ESE0	
ID Q9ESE0	PRELIMINARY;
ID Q9ESE0	PRT; 478 AA

DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Epidermal growth factor receptor related protein.  
GN ERFP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=GASTRO-DUODENAL MUCOSA;  
RA Yu Y., Koshier J.A., Majumdar A.P.N.;  
RT "Cloning of a novel EGFR-related peptide: A putative negative  
RT regulator of EGFR";  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF187818; AAC17037.2; -  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR002174; Furin-like.  
DR Pfam: PF00757; Furin-like; 1  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
KW Receptor.  
KW SEQUENCE 478 AA; 53233 MW; CF873A8376G519E5 CRC64;

Query Match	18.8%	Score 704.5	DB 11	Length 478
Best Local Similarity	34.5%	Pred. No. 2.4e-44		
Matches 180	Conservative 62	Mismatches 180	Indels 99	Gaps 18

QY 36 ARGAA--STVCTCTDMLKLAPAPENHLDMLRHLGYGCGVQGNELITLPTVASISFL 93  
 Db 20 AAGCALEEKAKVCGCTSNRLTGLTFEEDHFLSLQRYNNCEVLGNLTETVYQRNYDLSFL 79  
 QY 94 QDIEGYGVYLIAHNOYRVPFLQRLIVRGQLEEDNYALVALVDNGDPPLNTPPYGASP 155  
 Db 80 KTIQEVAGYFLIAHNYVERISSEDLQIIRGNALYENTYALISLN-----YETNR 1222  
 QY 154 GGLRELOLRSLTEFLKGGVYLQIQRNPOLCYODTILMKDI---EKKNNQLALTLIDTNRSR 209  
 Db 130 TGLRELPRLNRQLQELLIGAVRESNPDLICNDTIDWRIQIVONVEFSSNSMDL---GSHPS 165

```

0Y 210 ACHPE5PMCKGKSRCKMGESSEEDCOSTLTRRYVAGGGA - RKG9LPIDDCCHBOCAACCTPKH 266
Db 186 SCPKDPSCPMGSCMGGEENCKKLTJIKIAQOQSHRCKRGRSPSDCCCHNDCAACCTBPQK 243
0Y 269 SDCLACLPHNHSIGICELCPALVYNTDTPESMNPDEGRYFPGASCYTACPYNLTSDVG 328
Db 246 SDCLVCQKQFQDEATCKDTCPLMLXNPTTYMDVNPBEKYSFPGATCYAKNCPRANVTVDHG 300
0Y 329 SGAGMVHHRHRSSSTRSGGDLTLGLEPSEEARPSRLAPSEGASDVPEDGJGMAAK 388
Db 306 -----SCVRACGPD-----YVEVEDGIRK -CKKCDGPRCKVUNG -IGIEBK 346
0Y 389 GLOSLPTHDPSPLOAYSEDPTVPPLSETDGYVAPLTCSPDEYVNOQPDVQRPSPPEGR 448
Db 347 DTLST--NATNIKIFKXTAI-----SGDILLPLVA-----FKGDSFTTRPLDPRE-- 391
0Y 449 LPAARPAATLEBKTKTSLPGKNGV-----KDVAFGAVVNPPEL---TPQG--- 494
Db 392 -----LEIKTYVEITGSLLIQAMPENMTDLAAE-----FNLEIRGRTKOHOQFS 433
0Y 495 -----AAROPHPAPS-----PAEDNLYYMODP 519
Db 438 LAVGLNTLTSLEPOVPSLSMOAVTRPLRPLAONKRYSHDTGP 478

```

RESULT 13	
Q9BUD7	
ID Q9BUD7	PRELIMINARY;
331 AA	PRT;

DT 01-JUN-2001 (Tremblay, 17, Created)  
 DT 01-JUN-2001 (Tremblay, 17, Last sequence update)  
 DT 01-JUN-2002 (Tremblay, 21, Last annotation update)  
 DE Similar to v-erb-b2 avian erythroblastic leukemia viral oncogene  
 DE homolog 3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId:9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC002706; AA02706.1; -;  
 DR InterPro; IPR000494; EGFR\_L\_domain.  
 DR InterPro; IPR002174; Furin-like.  
 DR Pfam; PF00757; Furin-like; 1.  
 DR Pfam; PF01030; Recep\_L\_domain; 1.  
 DR SMART; SM00261; FU\_2.  
 SO SEQUENCE 331 AA; 36489 MW; 45E8EBE683FE7E8 CRC64;

Query Match	18.6%	Score	697	DB	4	Length	331
Best Local Similarity	44.6%	Pred	No. 5.3e-44				
Matches	139	Conservative	45	Mismatches	98	Indels	30
						Gaps	7

OY	16	IGFIFLFEFWIDRSVLAKELARGAA---SVOJGCTGDKKLYPASPENHMLMRYKVOGC	72
Db	10	LGLEF-----SLRGSVGVNSQAVCPETLNGLSVYTGDAEOMQYTLKYLIERC	56
OY	73	QVVOGNLELTYLPTNVASLSFLDIOEVOGYVLIAHNQVROPLORLRIVRSTOLEFNRYA	132
Db	57	EYVMGNLEIIVLGNHADSIFLOMIREVYGLVAMNNESTPLPLNLRVVRSTOYYDDKFA	116
OY	133	LAVLDNGDPLNNTTPVYGASPGGLRELOLRSVLTGLKGVLIOKRPOLCYOTDITLMDIF	199
- Db	117	IFVM-----LNYNT-----NSSHALROLRLTDLTEITSGGVYIEKNDKICHHDTIDMDIV	167
OY	193	HKNNQALATLIDTNRSRACHPCSPMGKSGRSWGESSEPCOSLTPTVCAGC--ARCCKPLP	251
Db	168	RDRD---AEIYVKDGNRSGPCPEHYCKG--KMGGSEDCOTLITITCAPQCNHGCFBPNP	223
OY	252	TDCCHEQCAACGCTGPKHSDCLACLHFNHSGICELHCPALVYVNTDFESMPNREGRTYFG	311

```

Db 224 NQCHDCCAGCGSPDQTDCAFCHFNDSGACVPCPQPLVYKNTLFOLEPNHTKYOG 283
QY 312 ASCVTACPYNYL 323
Db 284 GVCVASCPHFV 295

RESULT 14
QYH40
ID 09YH40 PRELIMINARY: PRT: 1165 AA.
AC 09YH40:
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Receptor tyrosine kinase proto-oncogene.
GN XMRK.
OS Xiphophorus xiphidium.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_Taxid=8086;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION;
RX MEDLINE=98241172; Pubmed=9582016;
RA Dimitrijevic N., Winkler C., Wellbrock C., Gomez A., Duschl J.,
RA Altschmidt J., Scharlt M.;
RT "Activation of the Xmrk proto-oncogene of Xiphophorus by
RT overexpression and mutational alterations.";
RL Oncogene 16:1681-1690(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION;
RA Scharlt M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U53471; AAD10500.2;
DR HSSP: P11362; 1FGK.
DR InterPro: IPR000345; Cyt_c_heme_bind.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_kinase; 1.
DR SMART: SM00261; FUF; 3.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 2.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; kinase; transferase; tyrosine-protein kinase.
SQ SEQUENCE 1165 AA; 129614 MW; 7E7EE38D8771A74E CRC64;

Query Match 17.8%; Score 664.5; DB 13; Length 1165;
Best Local Similarity 41.4%; Pred. No. 7.4e-41;
Matches 143; Conservative 46; Mismatches 129; Indels 27; Gaps 9;
QY 43 QVCTGDMKLRPASPETHLDMRLHLYOGGVQVGNLELYLTPTNLSLFLQDIOEVQY 102
Db 28 KYCGTSGNMTM--LDNHYLKMKKMYSGCNVYLENLEITYTQENDLSLQSIQEGY 84
QY 103 VLIANQVQVPLQRLRIYRGITOLFEDNVALAVLDGDPPLNTPTVGTGASPGGLRELQRL 162
Db 85 VLIANQVSTPLVNLRLRNGONLYEGNFTLVMSNYOK--NPSSP--DYVQVGLKQLOLS 141
QY 163 SLTEILKGVLLQIRNPQCLQDTILMKDIFHKNNQLATLILIDIRBRACHPCSPMKGSR 222

```

```

Db 142 NLTEILSGKGVSHNPLLCVNETIMMWDIVDKTSNPTMNLIPHAFFQCCRCPCGNGS 201
QY 223 CMGSESSDCOSLIPRTVAGGC-ARCKGPLPDDCCAGCAGTGPCKHSDCLACHFNHSG 281
Db 202 CWAPGPGCHCKFTKLCAECCNRRCRGPRIIDCCNEHAGAGCTPRATDCLACDFDNDG 261
QY 282 ICELHCPALVYNTDTPESNPNEGRTYFGASCVTACPYNYLSTDVSGAGVNHHRHS 341
Db 262 TCKDTCPPKIYDIVSHQVVDNPNIKYTFGACVKECPSNVYTE-----GACV----RS 312
QY 342 SSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVVDGLKCA 386
Db 313 CSA-----GMLEVDENGKRS-CKPCDGVCPKWDG-IGIGS 346

RESULT 15
QYGL44
ID 09GL44 PRELIMINARY: PRT: 144 AA.
AC 09GL44:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor.
GN GM-CSF.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_Taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Hutchinson K.L., Villinger F., Miranda M.E., Ksiazek T.G.,
RA Peters C.J., Rollin P.E.;
RT "Multiplex analysis of cytokines in the sera of cynomolgus macaques
RT naturally infected with Ebola (Reston).";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF007376; AAG16626.1;
DR HSSP: P0141; 2GMF.
DR InterPro: IPR000773; GM-CSF.
DR Pfam: PF01109; GM-CSF; 1.
DR PRINTS: PR00693; GMCSFACTOR.
DR ProDom: PD007349; GM-CSF; 1.
DR SMART: SM00040; CSF2; 1.
DR PROSITE: PS00702; GM-CSF; 1.
FT VARIANT 60 60 V -> I.
SQ SEQUENCE 144 AA; 16177 MW; 7D5F381DA2FC832F CRC64;

Query Match 17.5%; Score 654; DB 6; Length 144;
Best Local Similarity 94.5%; Pred. No. 2.8e-41;
Matches 121; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 548 AAPARSPSPQTPWEHNAIOEARRLNLSRDTAAENKTEVEYSEMFDLQEPCTQTRL 607
Db 17 SARARSSPTQWEHNAIOEARRLNLSRDTAAENKTEVEYSEMFDLQEPSCQTRL 76
QY 608 ELIKGILGSLTLKGPFLTMASHYKQCPPTPETSQATITTFESKKNLQPLVYIP 667
Db 77 ELKKGILGSLTLKGPFLTMASHYKQCPPTPETSQATITTFESKKNLQPLVYIP 136
QY 668 DCWEPVQ 675
Db 137 DCWEPVQ 144

```

Search completed: April 28, 2003, 13:41:31  
Job time : 52.1538 secs



GenCore version 5.1.4.P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:23 ; Search time 14.4231 Seconds

(Without alignments)  
1984.228 Million cell updates/sec

Title: US-09-821-883-2  
Perfect score: 3739  
Sequence: 1 MRAAPLLARASLSIGFLF.....EPVQEGAPPPAAHHHHH 690

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	64.3	1255	1	ERB2_HUMAN
2	1947.5	52.1	1257	1	ERB2_RAT
3	1935	51.8	1254	1	ERB2_MESAU
4	755.5	20.2	703	1	EGFR_CHICK
5	721	19.3	1210	1	EGFR_HUMAN
6	721	19.3	1308	1	ERB4_HUMAN
7	717.5	19.2	1210	1	EGFR_MOUSE
8	717	19.2	1308	1	ERB4_RAT
9	697	18.6	1342	1	ERB3_HUMAN
10	674	18.0	144	1	CSF2_HUMAN
11	668	17.9	1339	1	ERB3_RAT
12	655.5	17.5	1167	1	XMRK_XIPMA
13	549	14.7	144	1	CSF2_SHEEP
14	536	14.3	1426	1	EGFR_DROME
15	529	14.1	144	1	CSF2_CEREL
16	488	13.1	144	1	CSF2_PIG
17	481.5	12.9	143	1	CSF2_BOVIN
18	476.5	12.7	144	1	CSF2_CANFA
19	452	12.1	144	1	CSF2_FELCA
20	441	11.6	127	1	CSF2_RAT
21	435.5	11.6	140	1	CSF2_CAVPO
22	404.5	10.8	1323	1	LT23_CAEEL
23	367	9.8	141	1	CSF2_MOUSE
24	317.5	8.5	1363	1	ILPR_BRALA
25	293	7.8	245	1	ERB2_MOUSE
26	273.5	7.3	1477	1	HTK7_HYDAT
27	270.5	7.2	2146	1	INSR_DROME
28	257.5	6.9	1382	1	INSR_HUMAN
29	252.5	6.8	1372	1	INSR_MOUSE
30	252	6.7	1607	1	MIPR_LYMST
31	251.5	6.7	1300	1	IRR_MOUSE
32	249	6.7	1383	1	INSR_RAT
33	248.5	6.6	1297	1	IRR_HUMAN

34	246.5	6.6	1300	1	IRR_CAVPO	P14617 cavia porce
35	246	6.6	1367	1	IGIR_HUMAN	P6063 ratius norv
36	235.5	6.3	1370	1	IGIR_RAT	P60755 mus musculu
37	233.5	6.2	1373	1	IGIR_MOUSE	O64718 ratius norv
38	228.5	6.1	581	1	IRR_RAT	O64718 ratius norv
39	207	5.5	1390	1	INSR_AEDAE	O64718 ratius norv
40	178	4.8	2715	1	TRX2_HUMAN	O64718 ratius norv
41	168	4.5	386	1	PPAP_HUMAN	O64718 ratius norv
42	166	4.4	707	1	PPAP_MOUSE	O64718 ratius norv
43	162	4.3	1321	1	IRRS2_MOUSE	O64718 ratius norv
44	160	4.3	830	1	SREC_HUMAN	O64718 ratius norv
45	157.5	4.2	605	1	BRL1_EBV	P03209 epstein-bar

## ALIGNMENTS

RESULT 1	ERB2_HUMAN	STANDARD:	PRT: 1255 AA.
ID	ERB2_HUMAN		
AC	P04626:		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)		
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (tyrosine kinase-type cell		
DE	surface receptor HER2) (MLN 19).		
GN	ERBB2 OR HER2 OR NCL OR NEU.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID:9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=66118663; PubMed=3003577;		
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,		
RT	"Similarity of protein encoded by the human c-erbB-2 gene to		
RT	epidermal growth factor receptor.";		
RL	Nature 319:230-234(1986).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86016729; PubMed=2999974;		
RA	Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,		
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,		
RA	Francine U., Levinson A., Ulrich A.,		
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor		
RT	shares chromosomal location with neu oncogene.";		
RL	Science 230:1132-1139(1985).		
RN	[3]		
RP	SEQUENCE OF 737-1031 FROM N.A.		
RX	MEDLINE=86016729; PubMed=2999974;		
RA	Semba K., Kamata N., Toyoshima K., Yamamoto T.,		
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the		
RT	c-erbB-1/epidermal growth factor receptor gene and is amplified in a		
RT	human salivary gland adenocarcinoma.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).		
RN	[4]		
RP	VARIANTS VAL-654 AND VAL-655.		
RX	MEDLINE=93194196; PubMed=8095488;		
RA	Ehman A., Low J., Wallace R.B., Wu A.M.,		
RT	"Characterization of a new allele of the human ERBB2 gene by allele-		
RT	specific competition hybridization.";		
RL	Genomics 15:426-429(1993).		
CC	-1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,		
CC	ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A		
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-		
CC	ALPHA AND AMPHIREGULIN.		
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein		
CC	tyrosine phosphate.		
CC	-1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS		
CC	(POTENTIAL).		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.		

CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
 CC RESIDUES (BY SIMILARITY).  
 CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN  
 CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY  
 CC OF 0.782; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;  
 CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: M11767; AAA35808.1; JOINED.  
 DR EMBL: M11761; AAA35808.1; JOINED.  
 DR EMBL: M11762; AAA35808.1; JOINED.  
 DR EMBL: M11763; AAA35808.1; JOINED.  
 DR EMBL: M11764; AAA35808.1; JOINED.  
 DR EMBL: M11765; AAA35808.1; JOINED.  
 DR EMBL: M11766; AAA35808.1; JOINED.  
 DR EMBL: M11730; AAA75493.1; JOINED.  
 DR EMBL: M12036; AAA35978.1; JOINED.  
 DR EMBL: X03363; CAA27060.1; JOINED.  
 DR PIR: A25491; A25491.  
 DR PIR: A24571; A24571.  
 DR HSSP: P11362; 1FGK.  
 DR Gene: HGNC:3430; ERBB2.  
 DR MIM: 164870; ERBB2.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; kinase.1.  
 DR Pfam: PF00757; Furin-like.1.  
 DR Pfam: PF01030; Recep\_L\_domain.2.  
 DR Pfam: PF02757; YLP.2.  
 DR ProDom: PD000001; Euk\_kinase.1.  
 DR SMART: SM00261; FU.3.  
 DR SMART: SM00219; TyrKC.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR.1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM.1.  
 DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;  
 KW transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Polymorphism.  
 FT CHAIN 1 21 POTENTIAL.  
 FT DOMAIN 22 1235 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.  
 FT TRANSMEM 22 652 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 720 987 PROTEIN KINASE  
 FT NP\_BIND 726 734 ATP (BY SIMILARITY).  
 FT BINDING 733 733 ATP (BY SIMILARITY).  
 FT ACT\_SITE 845 845 BY SIMILARITY.  
 FT DISULFID 199 204 BY SIMILARITY.  
 FT DISULFID 199 212 BY SIMILARITY.  
 FT DISULFID 220 227 BY SIMILARITY.  
 FT DISULFID 224 235 BY SIMILARITY.  
 FT DISULFID 236 244 BY SIMILARITY.  
 FT DISULFID 240 252 BY SIMILARITY.  
 FT DISULFID 255 264 BY SIMILARITY.  
 FT DISULFID 268 295 BY SIMILARITY.  
 FT DISULFID 299 311 BY SIMILARITY.  
 FT DISULFID 315 331 BY SIMILARITY.  
 FT DISULFID 334 338 BY SIMILARITY.  
 FT DISULFID 511 520 BY SIMILARITY.  
 FT DISULFID 515 528 BY SIMILARITY.  
 FT DISULFID 531 540 BY SIMILARITY.  
 FT DISULFID 544 560 BY SIMILARITY.

FT DISULFID 563 576 BY SIMILARITY.  
 FT DISULFID 567 584 BY SIMILARITY.  
 FT DISULFID 587 596 BY SIMILARITY.  
 FT DISULFID 600 623 BY SIMILARITY.  
 FT DISULFID 626 634 BY SIMILARITY.  
 FT DISULFID 630 642 BY SIMILARITY.  
 FT MOD\_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 654 654 I -> V.  
 FT VARIANT 654 654 /FTID-VAR\_004077.  
 FT VARIANT 655 655 I -> V.  
 FT CONFLICT 1170 1170 /FTID-VAR\_004078.  
 FT CONFLICT 1170 1170 P -> A (IN REF. 2).  
 FT SEQUENCE 1255 AA; 137909 MM; 39E9DFDA04DCE962 CRC64;  
 SQ  
 Query Match 64.3%; Score 2405.5; DB 1; Length 1255;  
 Best Local Similarity 41.1%; Pred. No. 2.2e-124;  
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;  
 QY 35 LANGAASVQVCTGDMKRLPASPETHLDMRLHYQGCQVQVGNLELYLPTNASLSFLQ 94  
 DB 16 LPFGAASVQVCTGDMKRLPASPETHLDMRLHYQGCQVQVGNLELYLPTNASLSFLQ 75  
 QY 95 DIOEVQGYVLIANQVQVQLRLRYRGTOLEEDNVALVLDNGDPLNNTPTTGASPG 154  
 DB 76 DIOEVQGYVLIANQVQVQLRLRYRGTOLEEDNVALVLDNGDPLNNTPTTGASPG 135  
 QY 155 GLRELQRLSTLEIKGVLVLRNPOLCYOTILMKDIFHNNQALATLITNRSRACHPC 214  
 DB 136 GLRELQRLSTLEIKGVLVLRNPOLCYOTILMKDIFHNNQALATLITNRSRACHPC 195  
 QY 215 SPMCKSGRCWGESEEDCSLTRVCAGGACGKRLPTDCHEGACAGTGPKHSDCLAC 274  
 DB 196 SPMCKSGRCWGESEEDCSLTRVCAGGACGKRLPTDCHEGACAGTGPKHSDCLAC 255  
 QY 275 LHFHSGICELHCPALVTYNTDFEPMNPBGRRTFASCVTACPVYLSSTDGSS----- 329  
 DB 256 LHFHSGICELHCPALVTYNTDFEPMNPBGRRTFASCVTACPVYLSSTDGSSCTLYC 315  
 QY 330 ----- 329  
 DB 316 PLNHQVTAEDGTORCEKSKPCARVCYGLGMEHLREAVRANTSANIOEFAGCKRIEGLA 375  
 QY 330 ----- 329  
 DB 376 FLPESEFGDPASNTAPLPQEPQOLFETLEITGYLISAMPDLSPLDSVFQNLQVINGRI 435  
 QY 330 ----- 329  
 DB 436 LHNQAVSLTLQGLISWGLRSRELGSGLALHHNTHLCFVHTVPWDQLFRNPQALLH 495  
 QY 330 ----- 329  
 DB 496 TANRPEDECVGEGLAGHQLCARGHGNGPPTQCVCNCSQFLRGQCEVCECHVLOGLPREY 555  
 QY 330 ----- 329  
 DB 556 NARHCLPCHPEQOPQNSVTCFGEADQCAVCAHYKDPFCVACRPSGVRPDLSTYPMWK 615  
 QY 330 ----- 329  
 DB 616 FPDEGACQPCPINCCHSCVDLDKCGCPABORASPLTISAVVGLLVVGVFGILI 675  
 QY 330 ----- 329  
 DB 676 KRQQRKIRKTYMRRLDQETELVEPLTPSGAMPNQAQRIKETELRKVKVVGSGAGCTGY 735



```

FT DISULFID 569 586 BY SIMILARITY.
FT DISULFID 589 598 BY SIMILARITY.
FT DISULFID 602 625 BY SIMILARITY.
FT DISULFID 628 636 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT MOD_RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).
SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;

Query Match 52.1%; Score 1947.5; DB 1; Length 1257;
Best Local Similarity 34.2%; Pred. No. 2.4e-99;
Matches 433; Conservative 21; Mismatches 70; Indels 741; Gaps 3;

OY 11 AASLSIGFLFLFWMIDRSYLAKELARGAASVQVCTGTDMKRLPASPETHLDMRLHYO 70
DB 4 AAMCRMGFLAL-----LPPGIAGTGVCTGTDMKRLPASPETHLDMRLHYO 51
OY 71 GCQVVGNELELYLPTNASELFLQDIOEVGYLIAHNOVROPLORLRVGTQLPEDN 130
DB 52 GCQVVGNELELYLPTNASELFLQDIOEVGYLIAHNOVROPLORLRVGTQLPEDN 111
OY 131 YALAVDNDGPLNNTTTPVT-GASPGGLRELOLSLREILKGVLIORNPOLCYODTILMK 189
DB 112 YALAVDNDNDPDONVASTPGRTPBGLRELOLSLREILKGVLIORNPOLCYODVLMK 171
OY 190 DLFHKNQALTLIDTNRSPACHPCSPMCKSGRCMGESSEDCSLTRVYACAGCARCKGP 249
DB 172 DYFRKNNQALPVIDIDTNRSPACHPCSPMCKSGRCMGESSEDCQILNGTICTSCACRKR 231
OY 250 LPTDCHECCACAGCTGPKHSDDLACLAHFNHSGICELHCPALVYNNNDTPESMNPREGRT 309
DB 232 LPTDCHECCACAGCTGPKHSDDLACLAHFNHSGICELHCPALVYNNNDTPESMNPREGRT 291
OY 310 FGASCTACPYNYLSTDVGS----- 329
DB 292 FGASCTTCPCYNYLSTEVSGCTLYCPRNNOEYVADGTCRCKGSPCARVCTGLGMEHL 351
OY 330 ----- 329
DB 352 RGARATSDNVQFEDGCKKIFGSLAFLPESFDGDPSSGIALPRLPEQLQVETLEITGYL 411
OY 330 ----- 329
DB 412 YISAMPDSLRLDSVFONLRIINGRIIHDGAYSLTLOGLIHSGLRLSRELGSGLALIR 471
OY 330 ----- 329
DB 472 NMLICFVHYVPMQDLFRNPHQALLHSGNRPEEDLCVSSGLVCSLCAHGHGMPGPTQCV 531
OY 330 ----- 329
DB 532 NCSHFLRGOECVECHVMKGLPREVYSDKRLCPHCECOPONSSETCFGEADCAACAH 591
OY 330 ----- 329
DB 592 YNDSSCAVRCPSGVKPLSYPMWKYPDEEGICOPINCHSCVLDLBERGCPABORAS 651
OY 330 ----- 329
DB 652 PVTFIATVVGVLFLILLVYVVGILKRRRQKIRKTYMRLLDETLEVEPLPSGAMPNO 711
OY 330 ----- 329
DB 712 AQMRILKETELRKVKVLSGAGCTGYKGIWPDGENVKIPVAIKVLRNTPSKANKEILD 771
OY 330 ----- 329

```

```

DB 772 EAYVAVGSPYVSRLLGICITSTVOLVQIMPYGCLLDHVHREHGRGLSQDILLNCVOY 831
OY 330 ----- 329
DB 832 AKCMSTLEDVRLVHRLAARNLVKSPNHVKTIDFGLARLLDIDETRYHADGKVPFKMM 891
OY 330 ----- 329
DB 892 ALESLIRRTHQSDVWSYGVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPICTI 951
OY 330 ----- 329
DB 952 DVMYIMKCMWIDSECRPRERELVSFRRMARDPORFVYIQNEDLGSSPSPDSTFYRSL 1011
OY 330 ----- 361
DB 1012 EDDMDGLVDAEYLYLPQGFSPDPTPGTSTAHRHNSSTRSGGELTLGLPSEEG 1071
OY 362 APRSPLAPSEGAGSDVFDGDLGMAKGLQSLPTPHDSPLOQRYSEDPVPLPSETDGYVA 421
DB 1072 PRPSPLAPSEGAGSDVFDGDLGMAKGLQSLPHDLSPLQRYSEDPVPLPSETDGYVA 1131
OY 422 PLTCSQPEYVNOVDYRPPSPREGPLPAARAGATLERAKTLSGKNGVAVDAFAFG 481
DB 1132 PLTCSQPEYVNOSEVQPPPLPEGPPLPRPAGATLERAKTLSGKNGVAVDAFAFG 1191
OY 482 AVENPEYLPQGGAPAPPPAFSPAFDNLVYWDOPPERGAPSTFKGTFAENPEYL 541
DB 1192 AVENPEYLPREGTASPRHSPAFDNLVYWDONSSQGGPPNFECTPAENPEYL 1251
OY 542 GLDVP 546
DB 1252 GLDVP 1256

RESULT 3
ID ERB2_MESAU STANDARD; PRT; 1254 AA.
AC 060553;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (neu proto-oncogene) (C-erbB-2).
GN ERBB2 OR NEU.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Nerve;
RX MEDLINE=94193007; PubMed=7908275;
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,
RA Yamazaki Y., Ishikawa T.;
RT "Cloning and activation of the Syrian hamster neu proto-oncogene."
RL Gene 140:251-255(1994).
CC - FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC - SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES.
CC - SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```



QY 378 FDGDLGGAAGKGLQSLPTHPSPLOQRYSEDPVPLPSETDGYVAPLTCSPQPEVYNPDV 437  
 DB 1086 FEGELGAGARKGQSPISRLPSQRYSEDPPLPLPTETDGYVAPLACSPQPEVYNPDV 1145  
 OY 438 RPOPSRRECPPLPARAGATLIERAKTLSPGKNGVADVAFGAVANPEYLPQGGAP 497  
 DB 1146 RPOPLPEGLPLPVYRAGATLIERAKTLSPGKNGVADVAFGAVANPEYLPQGGAS 1205  
 OY 498 QPHPPAFSPAFNLYYWDODPPERGAPSPSTFKGTPTAENPEYLGIDVP 546  
 DB 1206 QPH-PPALCPAFNLYYWDODPPERGAPSPSTFKGTPTAENPEYLGIDVP 1253

RESULT 4  
 EGF\_R\_CHICK STANDARD; PRT; 703 AA.  
 AC P1387;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)  
 GN (fragment).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88261272; PubMed=3260329;  
 RA Lax I., Johnson A., Howk R., Sap J., Bellot F., Winkler M.,  
 RA Ullrich A., Vennstrom B., Schlessinger J., Glavol D.;  
 RT "Chicken epidermal growth factor (EGF) receptor: cDNA cloning,  
 RT expression in mouse cells, and differential binding of EGF and  
 RT transforming growth factor alpha."  
 RT Mol. Cell. Biol. 8:1970-1978(1988).  
 RL

CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,  
 CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND  
 CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to  
 CC dimerization, internalization of the EGF-receptor complex,  
 CC induction of the tyrosine kinase activity, stimulation of cell DNA  
 CC synthesis, and cell proliferation.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M20386; AAA48760.1;  
 DR InterPro: IPR000494; EGF\_R\_L\_domain.  
 DR InterPro: IPR000719; Euk.pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr.pkinase.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR SMART: SM00261; FU; 4.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; PARTIAL.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; PARTIAL.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; PARTIAL.  
 DR Transmembrane; Glycoprotein; Receptor; Signal; Transferase;  
 KW Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
 FT SIGNAL 1 30  
 FT CHAIN 31 >703  
 FT DOMAIN 31 654  
 FT EPIDERMAL GROWTH FACTOR RECEPTOR.  
 FT EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 655 667  
 FT DOMAIN 668 >703  
 FT DISULFID 197 206  
 FT DISULFID 201 214  
 FT DISULFID 222 230  
 FT DISULFID 226 238  
 FT DISULFID 239 247  
 FT DISULFID 243 255  
 FT DISULFID 258 267  
 FT DISULFID 271 298  
 FT DISULFID 302 314  
 FT DISULFID 318 333  
 FT DISULFID 336 340  
 FT DISULFID 513 522  
 FT DISULFID 517 530  
 FT DISULFID 533 542  
 FT DISULFID 546 562  
 FT DISULFID 565 581  
 FT DISULFID 569 589  
 FT DISULFID 592 601  
 FT DISULFID 605 627  
 FT DISULFID 630 638  
 FT DISULFID 634 646  
 FT CARBOHYD 134 134  
 FT CARBOHYD 190 190  
 FT CARBOHYD 200 200  
 FT CARBOHYD 359 359  
 FT CARBOHYD 368 368  
 FT CARBOHYD 420 420  
 FT CARBOHYD 573 573  
 FT CARBOHYD 578 578  
 FT CARBOHYD 613 613  
 FT CARBOHYD 633 633  
 FT CARBOHYD 648 648  
 FT NON TER 703 703  
 SQ SEQUENCE 703 AA; 77427 MW; AFD2E1B735A690 CRC64;

Query Match 20.2%; Score 755.5; DB 1; Length 703;  
 Best Local Similarity 41.9%; Pred. No. 1.7e-34;  
 Matches 161; Conservative 57; Mismatches 119; Indels 47; Gaps 10;

OY 28 RSVLAKELARCA-----STQVCTGDMKRLRASPETHLMDRLHYOG 71  
 DB 4 RSPLSASGPRGAALVLLLLGVALCSAVEERKVCQTNNKLTQLGVEDHFTSLQRYNN 63  
 OY 72 QOVQNGMLETLYLPTNASLFLDIOQVGYVLAHQVROVPLRLRYVGTQLEFDNY 131  
 DB 64 CEVLSMLETITYVENNDLFLKTIQVAGVILALMMVDVIPLENIQIIRGNVLYNSF 123  
 OY 132 ALAVLNGDPLNNTPTVYGASPGGLRELQRLSTELLKGVLLQIRNPOLCYODTILMKDI 191  
 DB 124 ALAVLSNYH-MNKTQ-----GLRELPMKRLEIILNGGVKIKSNBNKLCNMVDLWMDI 174  
 OY 192 FHKNOALFLID-TNRSRACHPCSPMKGRSGESEDCCQSLTRVCAAGCA-RCKP 249  
 DB 175 IDTSRK-PLTVLDFASMLSSCPCHPCTEDHCGAGEQNCQTLTKVICAQCSGRGK 233  
 OY 250 LPTDCHEQCAAGCTGPKHSDCLACLHFNHSGICELCPALVYNTDFTCSMPREGRYT 309  
 DB 234 VPSCCHNQCAGCTGPRSDCLACRKFRDQATCKDPCPLVLYNPTTYMDVAPBEKYS 293  
 OY 310 FGASCVTACPYNYLSTDVSGAGAGMVHRRSSSTRSGGDDTLGLEPSEDEAPRPLAP 369  
 DB 294 FGATCVRECPHNHYVTVDHGSVCV-----RSCNTDT-----YVEENGVRK-CKK 335  
 OY 370 SEGAGSDVPEQDGLGMAKGLQSL 393  
 DB 336 CDGLCSKVCNG-IGIGELKGLIST 358

RESULT 5  
 EGF\_R\_HUMAN STANDARD; PRT; 1210 AA.  
 ID EGF\_R\_HUMAN

AC P00533; P06268; Q14225; Q9UMD7; Q9UMD8; Q9UMG5; Q92795; Q00732;  
AC O00688; Q9BZ52; Q9BZC9; Q9BZC9; Q9BZC9;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor  
DE protein-tyrosine kinase ErbB-1).  
GN EGFR OR ERBB1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE-84219729; PubMed-6328312;  
RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,  
RA Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,  
RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;  
RT "Human epidermal growth factor receptor cDNA sequence and aberrant  
RT expression of the amplified gene in A431 epidermoid carcinoma cells.";  
RL Nature 309:418-425(1984).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX TISSUE-Placenta;  
RC MEDLINE-95382957; PubMed-7654368;  
RA Ilekis J.V., Stark B.C., Scoccia B.;  
RT "Possible role of variant RNA transcripts in the regulation of  
RT epidermal growth factor receptor expression in human placenta.";  
RL Mol. Reprod. Dev. 41:149-156(1995).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX TISSUE-Placenta;  
RC MEDLINE-97078686; PubMed-8918811;  
RA Reltter J.L., Maible N.J.;  
RT "A 1.8 kb alternative transcript from the human epidermal growth  
RT factor receptor gene encodes a truncated form of the receptor.";  
RL Nucleic Acids Res. 24:4050-4056(1996).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX TISSUE-Placenta;  
RC MEDLINE-97256547; PubMed-9103388;  
RA Ilekis J.V., Garfili J., Niederberger C., Scoccia B.;  
RT "Expression of a truncated epidermal growth factor receptor-like  
RT protein (TEGFR) in ovarian cancer.";  
RL Gynecol. Oncol. 65:36-41(1997).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
RX TISSUE-Placenta;  
RC MEDLINE-21100872; PubMed-11161793;  
RA Reltter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,  
RA Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampfand A.L.,  
RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,  
RA Maible N.J.;  
RT "Comparative genomic sequence analysis and isolation of human and  
RT mouse alternative EGFR transcripts encoding truncated receptor  
RT isoforms.";  
RL Genomics 71:1-20(2001).  
RN [6]  
RP SEQUENCE OF 575-687 FROM N.A.  
RA Reltter J.L., Threadgill D.W., Danielsen A.J., Schehl C.M.,  
RA Lampfand A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,  
RA Maible N.J.;  
RT "Human and mouse alternative EGFR transcripts encoding only the  
RT extracellular domain of the receptor.";  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE OF 713-924 FROM N.A.  
RX MEDLINE-84196372; PubMed-6326261;  
RA Lin C.R., Chen W.S., Krulger W., Stolarsky L.S., Weber W.,  
RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;  
RT "Expression cloning of human EGF receptor complementary DNA: gene  
RT amplification and three related messenger RNA products in A431  
RT cells.";  
RL Science 224:843-848(1984).  
RN [8]  
RP SEQUENCE OF 150-962 FROM N.A.  
RX MEDLINE-84245835; PubMed-6330563;  
RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,  
RA Roe B.A., Merlino G.T., Pastan I.;  
RT "Human epidermal growth factor receptor cDNA is homologous to a  
RT variety of RNAs overproduced in A431 carcinoma cells.";  
RL Nature 309:806-810(1984).  
RN [9]  
RP SEQUENCE OF 1028-1210 FROM N.A.  
RX MEDLINE-85046483; PubMed-6093780;  
RA Slimen F.A., Gope M.L., Schultz T.Z., Wright D.A., Carpenter G.,  
RA O'Malley B.W.;  
RT "Isolation of an evolutionarily conserved epidermal growth factor  
RT receptor cDNA from human A431 carcinoma cells.";  
RL Biochem. Biophys. Res. Commun. 124:125-132(1984).  
RN [10]  
RP SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE-88217333; PubMed-3329716;  
RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,  
RA Waterfield M.D.;  
RT "The human EGF receptor gene: structure of the 110 kb locus and  
RT identification of sequences regulating its transcription.";  
RN Oncogene Res. 1:375-396(1987).  
RN [11]  
RP SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE-91107677; PubMed-1988448;  
RA Haley J.D., Waterfield M.D.;  
RT "Contributory effects of de novo transcription and premature  
RT transcript termination in the regulation of human epidermal growth  
RT factor receptor proto-oncogene RNA synthesis.";  
RL J. Biol. Chem. 266:1746-1753(1991).  
RN [12]  
RP SEQUENCE OF 1-29 FROM N.A.  
RX MEDLINE-85270438; PubMed-2991899;  
RA Ishii S., Xu Y.H., Strutton R.H., Roe B.A., Merlino G.T., Pastan I.;  
RT "Characterization and sequence of the promoter region of the human  
RT epidermal growth factor receptor gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).  
RN [13]  
RP SEQUENCE OF 540.  
RA Kohda D.;  
RT Submitted (SEP-1997) to the SWISS-PROT data bank.  
RN [14]  
RP RECEPTOR ACTIVITY.  
RX MEDLINE-84191554; PubMed-6325948;  
RA Mroczkowski B., Mosig G., Cohen S.;  
RT "ATP-stimulated interaction between epidermal growth factor receptor  
RT and supercoiled DNA.";  
RL Nature 309:270-273(1984).  
RN [15]  
RP PHOSPHORYLATION.  
RX MEDLINE-89278137; PubMed-2543678;  
RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honeyger A.M.,  
RA Hawk R., Givol D., Ullrich A., Schlessinger J.;  
RT "All autophosphorylation sites of epidermal growth factor (EGF)  
RT receptor and HER2/neu are located in their carboxyl-terminal tails.  
RT Identification of a novel site in EGF receptor.";  
RL J. Biol. Chem. 264:10667-10671(1989).  
RN [16]  
RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND  
RP ASN-528.  
RX MEDLINE-86396132; PubMed-8962717;  
RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;  
RT "Analysis of the glycosylation patterns of the extracellular domain of  
RT the epidermal growth factor receptor expressed in Chinese hamster  
RT ovary fibroblasts.";  
RL Growth Factors 13:121-132(1996).  
RN [17]  
RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND  
RP ASN-603.  
RX MEDLINE-20198209; PubMed-10731668;

```

RA Sato C, Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;
RT "Characterization of the N-oligosaccharides attached to the atypical
RT Asn-x-Cys sequence of recombinant human epidermal growth factor
RT receptor.";
RL J. Biochem. 127:65-72(2000).
RN [18]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=98225196; PubMed=9556602;
RA Abe Y., Otake M., Inagaki F., Lax I., Schlessinger J., Kohda D.;
RT "Disulfide bond structure of human epidermal growth factor receptor.";
RL J. Biol. Chem. 273:11150-11157(1998).
RN [19]
RP REVIEW.
RX MEDLINE=87297456; PubMed=3039909;
RA Carpenter G.;
RT "Receptors for epidermal growth factor and other polypeptide
RT mitogens.";
RL Annu. Rev. Biochem. 56:881-914(1987).
CC -1- FUNCTION: Receptor for EGF, but also for other members of the EGF
CC family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding
CC EGF-like growth factors, GP30 and vaccinia virus growth factor. Is
CC involved in the control of cell growth and differentiation.
CC -1- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is
CC secreted.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/
CC truncated isoform/TEGFR, p310 and 4; are produced by
CC alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also
CC expressed in ovarian cancers.
CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to
CC dimerization, internalization of the EGF-receptor complex,
CC induction of the tyrosine kinase activity, stimulation of cell DNA
CC synthesis, and cell proliferation.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL; X00588; CAA25240.1; -
DR EMBL; U95089; AAB53063.1; -
DR EMBL; U48722; AAC50802.1; -
DR EMBL; U48723; AAC50804.1; -
DR EMBL; U48724; AAC50796.1; -
DR EMBL; U48725; AAC50797.1; -
DR EMBL; U48726; AAC50798.1; -

```

```

Oy 275 LAFNMSGJELCELPALVYNTNTDTFSPMSPEBERYFFGASVCACVYNTLSTDTGSGAGAGM 334
Db 232 KRFREDAIOTKDCPPMLMLXNPTTYQMDVNPBEKYSFGATCVAKKCRNRYVYTHG----- 305
Oy 335 VHHRRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDJGMAKQLOSL 393
Db 306 -----SCVRACGAD-----SYEMEDGVKR-CKKCEGPCRKVCNG-IGIGEFKDSLST 351

RESULT 6
ERB4_HUMAN STANDARD: PRT; 1308 AA.
ID ID_1
AC 015303:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)
DE (p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).
GN ERB4 OR HER4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID:9606;
RN RN_1
RP SEQUENCE FROM N.A. (ISOFORM JM-A).
RC TISSUE-Breast carcinoma:
RX MEDLINE-93189574; PubMed-8383326;
RA Plozman G.D., Culouscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,
RA Foy L., Neubauer M.G., Shoyab M.;
RT "Ligand-specific activation of HER4/p180erbB4, a fourth member of the
RT epidermal growth factor receptor family.";
RT Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).
RN RN_2
RP SEQUENCE FROM N.A. (ISOFORM JM-A AND JM-B).
RC TISSUE-Fetal brain:
RX MEDLINE-97476287; PubMed-9334263;
RA Elenius K., Cortas G., Paul S., Choi C.J., Rlo C., Plozman G.D.,
RA Klagsbrun M.;
RT "A novel juxtamembrane domain isoform of HER4/erbB4, isoform-specific
RT tissue distribution and differential processing in response to
RT phorbol ester.";
RT J. Biol. Chem. 272:26761-26768(1997).
RL RL_1
J-1. FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND
NTRK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.
NOT ACTIVATED BY EGF, TGF- $\alpha$ , AND AMPHIRGULIN.
CC CC_1
-1 CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
tyrosine phosphate.
CC CC_2
-1 SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE ERBB
RECEPTORS (POTENTIAL).
CC CC_3
-1 SUBCELLULAR LOCATION: Type I membrane protein.
CC CC_4
-1 ALTERNATIVE PRODUCTS: 2 ISOFORMS. JM-A (SHOWN HERE) AND JM-B;
ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER
FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED
BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND
NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN
CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.
CC CC_5
-1 TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,
KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,
PITUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,
LUNG, SALIVARY GLAND, AND PANCREAS.
CC CC_6
-1 PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
RESIDUES.
CC CC_7
-1 SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC CC_8
-----
CC CC_9
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on ways
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```



CC EMBL: L07868; AAB59446.1; -  
 DR HSSP: P11362; IEGK.  
 DR Genew: HGNC:3432; ERBB4.  
 DR MIM: 600543; -  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR Pfam: PF02757; YLP; 2.  
 DR Prodom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; FU; 4.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KW Alternative splicing.  
 KW SIGNAL 1 25  
 FT CHAIN 1 25  
 FT DOMAIN 26 1308  
 FT TRANSMEM 652 675  
 FT DOMAIN 676 1308  
 FT DOMAIN 186 334  
 FT DOMAIN 496 633  
 FT DOMAIN 718 985  
 FT NP\_BIND 724 732  
 FT BINDING 751 751  
 FT ACT\_SITE 843 843  
 FT DISULFID 189 197  
 FT DISULFID 193 205  
 FT DISULFID 213 221  
 FT DISULFID 217 229  
 FT DISULFID 230 238  
 FT DISULFID 234 246  
 FT DISULFID 249 258  
 FT DISULFID 262 289  
 FT DISULFID 293 304  
 FT DISULFID 308 323  
 FT DISULFID 326 330  
 FT DISULFID 503 512  
 FT DISULFID 507 520  
 FT DISULFID 523 532  
 FT DISULFID 536 552  
 FT DISULFID 555 569  
 FT DISULFID 559 577  
 FT DISULFID 580 589  
 FT DISULFID 593 614  
 FT DISULFID 617 625  
 FT DISULFID 617 625  
 FT MOD\_RES 1162 1162  
 FT MOD\_RES 1188 1188  
 FT MOD\_RES 1258 1258  
 FT MOD\_RES 1284 1284  
 FT CARBOHYD 138 138  
 FT CARBOHYD 174 174  
 FT CARBOHYD 181 181  
 FT CARBOHYD 253 253  
 FT CARBOHYD 358 358  
 FT CARBOHYD 410 410  
 FT CARBOHYD 473 473  
 FT CARBOHYD 495 495  
 FT CARBOHYD 548 548  
 FT CARBOHYD 576 576  
 FT CARBOHYD 620 620  
 FT VANSPLIC 626 648  
 FT SEQUENCE 1308 AA: 146807 MW: 5E4AE80985D88761 CRC64:

Query Match 19.3%: Score 721; DB 1; Length 1308;  
 Best Local Similarity 46.9%; Pred. No. 2,66-32;  
 Matches 142; Conservative 42; Mismatches 107; Indels 12; Gaps 4;  
 QY 22 LFFWIDRSVLAKELARGASVQCTGTDMLRLPASPTHLDMLHLYQCGVYVGNLEI 81  
 DB 7 LVMVWSSLVAAGTV-QPDSQSVCAGTENKLSLSLSDLEQYRALRKYYENCEVVGNIET 65  
 QY 82 TYLPINASLFLQDIOEVQYVLAAHNOYROYPLQRLNIVKSTOLFEDENYALAVDNDP 141  
 DB 66 TSIEHNRDLSEFRSVREYTGVLVALNOFRYLPLEHRIIRKTKYEDRYALALFLNRRK 125  
 QY 142 LNTPTVVGASPGGRLREYOLSLTFEILKGVLIQNPOLCQYODTLUMDIFHKNNQLAT 201  
 DB 126 DGNF-----GLQELGLKNLTELINGVYVVDNKKFLCYADTIHMODIYRNPMSILT 176  
 QY 202 LIDTNRBRACHPCSPMCKSRMGSESSDCOSLTFRTVAGGC-ARCKGPLPTDCHEOCA 260  
 DB 177 LVSTNGSSGCGRHKSCGTG-RQMGPTENHCQTLRTVCAEQDGRGCPYVSDCCRBCA 235  
 QY 261 AGCTGPKHSDCLACHFNHSGICEALCPALVTYNTDTESMPNPEGRITFGASCYTACP 320  
 DB 236 GGCSPKPDTCDFACMNFDSGACVYQCPQTFVYNPPTFQLEHNFNAKYTGAFVCKCPH 295  
 QY 321 NYL 323  
 DB 296 NFV 298  
 RESULT 7  
 EGFR\_MOUSE STANDARD: PRT: 1210 AA.  
 ID EGFR\_MOUSE  
 AC 001279;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112).  
 GN EGFR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; TISSUE-Liver;  
 RX MEDLINE-93026370; PubMed-1408137;  
 RA Avivi A., Skorecki K., Yayan A., Givol D.;  
 RT "Promoter region of the murine fibroblast growth factor receptor 2  
 (bek/KGFR) gene.";  
 RL Oncogene 7:1957-1962(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; and CD-1; TISSUE-Liver, and Decidua;  
 RX MEDLINE-93126380; PubMed-7678348;  
 RA Parla B.C., Das S.K., Andrews G.K., Dey S.K.;  
 RT "Expression of the epidermal growth factor receptor gene is regulated  
 in mouse blastocysts during delayed implantation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BALB/c; TISSUE-Liver;  
 RX Hibbs M.L.;  
 RT Submitted (Apr-1994) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-B6/C3; TISSUE-Liver;  
 RX MEDLINE-94170986; PubMed-812525;  
 RA Luetteke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,  
 RA Jenkins N.A., Lee D.C.;  
 RT "The mouse waved-2 phenotype results from a point mutation in the EGF  
 receptor tyrosine kinase.";  
 RL Genes Dev. 8:399-413(1994).

RN [5] SEQUENCE OF 1-714 FROM N.A.  
 RP TISSUE-Brain;  
 RC MEDLINE-91232866; PubMed-2030916;  
 RA Avila A., Lax I., Ullrich A., Schlessinger J., Givol D., Morse B.;  
 RT "Comparison of EGF receptor sequences as a guide to study the ligand  
 binding site." Oncogene 6:673-676(1991).  
 RL (6)  
 RN SEQUENCE OF 969-1117 FROM N.A.  
 RP STRAIN-C3H;  
 RC Elstner D.P., Serrero G.;  
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,  
 AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND  
 VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to  
 dimerization, internalization of the EGF-receptor complex,  
 induction of the tyrosine kinase activity, stimulation of cell DNA  
 synthesis, and cell proliferation.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X78987; CAA55587.1; -  
 DR EMBL: U03425; AAA17899.1; -  
 DR EMBL: X59698; CAA42219.1; -  
 DR EMBL: L06864; AAA53029.1; -  
 DR EMBL: Z12608; CAA78249.1; -  
 DR HSSP: P11362; 1FGK.  
 DR MGD: MGI:95294; Egfr.  
 DR InterPro: IPR000494; EGFR\_L\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR Pfam: PF00757; Furin-like; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 2.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00261; FU; 3.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;  
 KM Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.  
 FT CHAIN 1 24  
 FT DOMAIN 25 1210  
 FT TRANSMEM 647 670  
 FT DOMAIN 671 1210  
 FT REPEAT 75 300  
 FT REPEAT 390 600  
 FT DOMAIN 1028 1071  
 FT DOMAIN 714 961  
 FT NP\_BIND 720 728  
 FT BINDING 747 747  
 FT ACT\_SITE 839 839  
 FT DISULFID 190 199  
 FT DISULFID 194 207  
 FT DISULFID 215 223  
 FT DISULFID 219 231  
 FT DISULFID 232 240  
 FT DISULFID 236 248  
 FT DISULFID 236 248

FT DISULFID 251 260 BY SIMILARITY.  
 FT DISULFID 264 291 BY SIMILARITY.  
 FT DISULFID 295 307 BY SIMILARITY.  
 FT DISULFID 311 326 BY SIMILARITY.  
 FT DISULFID 329 333 BY SIMILARITY.  
 FT DISULFID 506 515 BY SIMILARITY.  
 FT DISULFID 510 523 BY SIMILARITY.  
 FT DISULFID 526 535 BY SIMILARITY.  
 FT DISULFID 539 555 BY SIMILARITY.  
 FT DISULFID 558 571 BY SIMILARITY.  
 FT DISULFID 562 579 BY SIMILARITY.  
 FT DISULFID 582 591 BY SIMILARITY.  
 FT DISULFID 595 617 BY SIMILARITY.  
 FT DISULFID 620 628 BY SIMILARITY.  
 FT DISULFID 624 636 BY SIMILARITY.  
 FT DISULFID 640 646 BY SIMILARITY.  
 FT MOD\_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).  
 FT MOD\_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 1197 1197 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 128 128 (BY SIMILARITY).  
 FT CARBOHYD 175 175 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 196 196 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 413 413 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 444 444 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 528 528 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 568 568 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 603 603 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 623 623 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CONFLICT 19 19 C -> S (IN REF. 2).  
 FT CONFLICT 539 539 C -> W (IN REF. 5).  
 FT CONFLICT 991 991 L -> F (IN REF. 4).  
 FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).  
 SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DFD2F5 CRC64;  
 Query Match 19.2%; Score 717.5; DB 1; Length 1210;  
 Best Local Similarity 46.8%; Pred. No. 3; 7e-32;  
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;  
 QY 36 ARGA--STOVCCTGDKRLPASPETHLDMLKVCQVGVNLETLPTNASTFL 93  
 DB 20 AAGGALAEKRVCGSTSRKRLQLOTFEDHPLSLQRMVNCVGLNLEITYVQRNYSFL 79  
 QY 94 QDIQEVGVYLAHNOVQVPLQRLIVRGTQLEFEDYVALAVLDNGDPLNNTTPVTCASP 153  
 DB 80 KTIQEVAGVYLAHNVVERIPRLNLTIRGNALYENTYALALISN-----YGTNR 129  
 QY 154 GGLRELQRLSTELILNGVLIQRLNPOLCYODTILMKDI-----FKKNQALALTLIDNRSR 209  
 DB 130 TGLRELPMRLQELILGAVRFSSNNPILCNMDTIQWRDILVONVSMNSMOL---OSHPs 185  
 QY 210 ACHPCSPMGSCMGSSSDCSLFTVCAGCA--RCKGPLPTDCHEQCAAGCPKH 268  
 DB 186 SCRCPCSPSCNGSCMGSGEENCKKRLKIIICAOQCSHRCSRSPSDCHNCAGCIGPRE 245  
 QY 269 SDCLACLFHNSGICELHCPALVYNTDFESMPNPEGRTGFCASCVTACPVYLTVDG 328  
 DB 246 SDCLVCKQFQDEATCKDTCPLMLYNTTYQMDVNEGKYSFGATGVKCPRVYVYVDHG 305  
 QY 329 S 329  
 DB 306 S 306  
 RESULT 8  
 ERB4\_RAT  
 ID ERB4\_RAT STANDARD; PRT; 1308 AA.  
 AC Q62956; Q922N7;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).  
GN ERBB4 OR TYRO-2.  
OS Rattus norvegicus (Rat).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Heart;  
RX MEDLINE=98221155; PubMed=9553078;  
RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,  
RA Marchionni M.A., Kelly R.A.;  
RT "Neuregulins promote survival and growth of cardiac myocytes.  
RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult  
RT ventricular myocytes.";  
RL J. Biol. Chem. 273:10261-10269(1998).  
RN [2]  
RP SEQUENCE OF 848-901 FROM N.A.  
RC TISSUE-Sciatic nerve;  
RX MEDLINE=91222560; PubMed=2025425;  
RA Lai C., Lemke G.;  
RT "An extended family of protein-tyrosine kinase genes differentially  
RT expressed in the vertebrate nervous system.";  
RL Neuron 6:691-704(1991).  
RN [3]  
RP SEQUENCE OF 1031-1198 FROM N.A.  
RC STRAIN-Sprague-Dawley; TISSUE-Spinal cord;  
RX MEDLINE=97184212; PubMed=9030624;  
RA Carroll S.L., Miller M.L., Frohner P.W., Kim S.S., Corbett J.A.;  
RT "Expression of neuregulins and their putative receptors, ErbB2 and  
RT ErbB3, is induced during Wallerian degeneration.";  
RL J. Neurosci. 17:1642-1659(1997).  
CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULIN, NRG-  
CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND  
CC NTRK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.  
CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
CC tyrosine phosphate.  
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB  
CC RECEPTORS (POTENTIAL).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING  
CC NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS  
CC OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE  
CC RETICULAR NUCLEUS OF THE THALAMUS. VERY LOW LEVELS IN KIDNEY, AND  
CC HEART.  
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE  
CC RESIDUES (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF041838; AAD08899.1; -  
DR EMBL: U52531; AAC53051.1; -  
DR HSSP: P11362; IEGK.  
DR InterPro: IPR000494; EGFR\_L\_domain.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002174; Furin-like.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR InterPro: IPR004019; YLP\_motif.  
DR Pfam: PF00757; Furin-like; 1.  
DR Pfam: PF00069; Pkinase; 1.  
DR Pfam: PF01030; Recep\_L\_domain; 2.  
DR Pfam: PF02757; YLP; 2.  
DR PRINTS: PR00109; TYRKINASE.  
DR PRODOM: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00261; FU; 4.

DR SMART: SM00219; TYRK; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;  
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
FT SIGNAL 1 25  
FT CHAIN 26 1308  
FT DOMAIN 26 651  
FT TRANSMEM 652 675  
FT DOMAIN 676 1308  
FT DOMAIN 186 334  
FT DOMAIN 496 633  
FT DOMAIN 718 985  
FT NP\_BIND 724 732  
FT BINDING 751 751  
FT ACT\_SITE 843 843  
FT DISULFID 189 197  
FT DISULFID 193 205  
FT DISULFID 213 221  
FT DISULFID 217 229  
FT DISULFID 230 238  
FT DISULFID 234 246  
FT DISULFID 249 258  
FT DISULFID 262 289  
FT DISULFID 293 304  
FT DISULFID 308 323  
FT DISULFID 326 330  
FT DISULFID 503 512  
FT DISULFID 507 520  
FT DISULFID 523 532  
FT DISULFID 536 552  
FT DISULFID 555 569  
FT DISULFID 559 577  
FT DISULFID 580 589  
FT DISULFID 593 614  
FT DISULFID 617 625  
FT DISULFID 621 633  
FT MOD\_RES 1162 1162  
FT MOD\_RES 1188 1188  
FT MOD\_RES 1258 1258  
FT MOD\_RES 1284 1284  
FT CARBOHYD 138 138  
FT CARBOHYD 174 174  
FT CARBOHYD 253 253  
FT CARBOHYD 358 358  
FT CARBOHYD 410 410  
FT CARBOHYD 473 473  
FT CARBOHYD 495 495  
FT CARBOHYD 548 548  
FT CARBOHYD 576 576  
FT CARBOHYD 620 620  
FT CONFLICT 1062 1062  
FT CONFLICT 1080 1082  
FT SEQUENCE 1308 AA; 146957 MW; D944BB096A0B41 CRC64;  
Query Match 19.28; Score 717; DB 1; Length 1308;  
Best Local Similarity 47.28; Posit. No. 42e-33;  
Matches 142; Conservative 41; Mismatches 106; Indels 12; Gaps 4;  
QY 25 WLDKSVL-AKELARGASTOVCTGTDMLRLPASPEHLDMLRHLVYGCQVYVGNLELY 83  
DB 8 WVMGSLVAVARTVOPASQSVACAGTENKLSLSLDEQGYALRYKRYEYCEVVMGNETLS 67  
QY 84 LPTNASLFDLDDIOEVQGYVLAHNVQVYVQLRLRYRGQLFEDNRYALVLDNGDPLN 143  
DB 68 IEHNRDLSPFLRSIREYGVYVVALNQPFLPLNRIIRGKLYEDRYALALFIINRRDG 127  
QY 144 NTPVVGASPGGLRELQRLSLTEILKGVLFQRPOLCYOYTIAMKDFHNNQALATLI 203  
DB 128 NF-----GLQDLGKNTLETILNGVYVDQKFLCYADTITMODIVIRIPWISNMTLV 178  
QY 204 DTNRSRACHPCSPMCKGSRGCESSDEDCQSLRTVACGCC-ARCKGPLPTDCCHQCAAG 262



Query Match 18.6%; Score 697; DB 1; Length 1342;  
 Best Local Similarity 44.6%; Pred. No. 5.4e-31;  
 Matches 139; Conservative 45; Mismatches 98; Indels 30; Gaps 7;

16 LGFLEFLFMDRSLVAKELARGAA---STOYCTGTDMKRLPASPETHLMDLRHLQGC 72  
 10 LGFLF-----SLARGSEVNSQAVCPCTLNGLSVTGDAENQOYLYLYERC 56  
 73 QVVGNNLELTLPPTNASISFLDQDQVGYLIANQVROPLORLRVRTOLFEDNYA 132  
 57 EVVGNNEIYVLTGHADISFLQWREVGYLVAMNESTPLPLRLRVROTQYDGF 116  
 133 LAVDNGDPLNNTPVTGSPGLRELOLRSLTEILKGVLIQRPOLCQYDTIMKDI 192  
 117 IFW-----LVNT-----NSSHALRQLRLTOLTELSGVYIEKNDKLCMDTIDMR 167  
 193 HKNQOLATLTDITNSRACHPCSPCKGRCWGESSEDQSLTRTVACGGC-ARCKG 251  
 168 RDPR---AEIVVKDNGRCPCPCHEVCKG-RCWGPSEDCQRLTFTICAPQCNGH 223  
 252 TDCCHGCAACCTGPKHSDCLACHFNISGICELHCPALVYNDTFESMNPERRTF 311  
 224 NQCHDEACGSGSPQDTDFCAHRHFNDSGACVPRCPQPLVYNTLQLEBNPHTK 283  
 312 ASCVTACPYNTL 323  
 284 GCVVASCPHFNV 295

RESULT 10  
 CSF2\_HUMAN  
 ID CSF2\_HUMAN STANDARD; PRT; 144 AA.  
 AC P04141;  
 DT 01-NOV-1986 (Rel. 03, Last sequence update)  
 DT 01-NOV-1986 (Rel. 03, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)  
 DE (Colony-stimulating factor) (CSF) (Sargramostim) (Molgramostin).  
 GN CSF2 OR GMCSF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85242684; PubMed=3925454.  
 RA Lee F., Yokota T., Otsuka T., Gemmell L., Larson N., Luh J.,  
 RA Arai K.-I., Rennick D.;  
 RT "Isolation of cDNA for a human granulocyte-macrophage  
 RT colony-stimulating factor by functional expression in mammalian  
 RT cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4360-4364(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86205844; PubMed=3486413;  
 RA Kaushansky K., O'Hara P.J., Berkner K., Segal G.M., Hagen F.S.,  
 RA Adamson J.W.;  
 RT "Genomic cloning, characterization, and multilineage growth-promoting  
 RT activity of human granulocyte-macrophage colony-stimulating factor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3101-3105(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85298329; PubMed=3898082;  
 RA Cantrell M.A., Anderson D., Cerretti D.P., Price V., McCreghan K.,  
 RA Tuninski R.J., Mochizuki D.Y., Larsen A., Grabstein K., Cosman D.,  
 RT "Cloning, sequence, and expression of a human granulocyte/macrophage  
 RT colony-stimulating factor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6250-6254(1985).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85218749; PubMed=3923623;  
 RA Wong G.G., Wittek J.S., Temple P.A., Wilkens K.M., Leary A.C.,

RA Luxenberg D.P., Jones S.S., Brown E.L., Kay R.M., Orr E.C.,  
 RA Shoemaker C., Golde D.W., Kaufman R.J., Hewick R.M., Wang E.A.,  
 RA Clark S.C.;  
 RT "Human GM-CSF: molecular cloning of the complementary DNA and  
 RT purification of the natural and recombinant proteins.";  
 RL Science 228:810-815(1985).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86030234; PubMed=3876930;  
 RA Miyatake S., Otsuka T., Yokota T., Lee F., Arai K.-I.;  
 RT "Structure of the chromosomal gene for granulocyte-macrophage colony  
 RT stimulating factor: comparison of the mouse and human genes.";  
 RL EMBO J. 4:2561-2568(1985).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,  
 RA Davis C.A., Kader K., Miguel T., Pittluck S., Pollard M., Rojaski H.,  
 RA Sudramanian S., Martin C.H.;  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A. AND VARIANTS ILE-115 AND THR-117.  
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
 RA Nickerson D.A.;  
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP CARBOHYDRATE-LINKAGE SITES.  
 RX MEDLINE=92144609; PubMed=1737041;  
 RA Kaushansky K., Lopez J.A., Brown C.B.;  
 RT "Role of carbohydrate modification in the production and secretion of  
 RT human granulocyte macrophage colony-stimulating factor in genetically  
 RT engineered and normal mesenchymal cells.";  
 RL Biochemistry 31:1881-1886(1992).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE=92108420; PubMed=1837174.  
 RA Dederichs K., Boone T., Karpus P.A.;  
 RT "Novel fold and putative receptor binding site of  
 RT granulocyte-macrophage colony-stimulating factor.";  
 RL Science 254:1779-1782(1991).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=92235844; PubMed=1569568;  
 RA Walter M.R., Cook W.J., Fallick S.E., Nagabhushan T.L., Trotta P.P.,  
 RA Bugg C.E.;  
 RT "Three-dimensional structure of recombinant human granulocyte-  
 RT macrophage colony-stimulating factor.";  
 RL J. Mol. Biol. 224:1075-1085(1992).  
 CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION  
 CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING  
 CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- POLYMORPHISM: VARIANT ILE-117 MAY BE A RISK FACTOR FOR ATOPIC  
 CC ASTHMA.  
 CC -1- PHARMACEUTICAL: Available under the names Leukine (Immunex) and  
 CC Leucomax (Novartis). Used in myeloid reconstitution following bone  
 CC marrow transplant, bone marrow transplant engraftment failure or  
 CC delay, mobilization and following transplantation of autologous  
 CC peripheral blood progenitor cells, and following induction  
 CC chemotherapy in older adults with acute myelogenous leukemia.  
 CC -1- DATABASE: NME-Leukine; NOTE-Clinical information on Leukine;  
 CC WWW="http://www.imunex.com/patient/pa02el.html".  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M13207; AAA98768.1; -;  
 DR EMBL; M11734; AAA52122.1; -;  
 DR EMBL; M11220; AAA52578.1; -;

```

DR EMBL: X03021; CAA26822.1; -
DR EMBL: M10653; AAA52121.1; -
DR EMBL: AC004511; AAC08707.1; -
DR EMBL: AF373868; AAK51563.1; -
DR PIR: A01853; FOHUGM.
DR PIR: C24636; C24636.
DR PIR: A25169; A25169.
DR PDB: 1CSG; 31-JAN-94.
DR PDB: 2GMF; 08-NOV-96.
DR Genew: HGNC:2434; CSF2.
DR MIM: 138960; -.
DR InterPro: IPR000773; GM_CSF.
DR Pfam: PF01109; GM_CSF.1.
DR PRINTS: PR00693; GMCSFACOR.
DR PRODOM: PD007349; GM_CSF.1.
DR SMART: SM00040; CSF2; 1.
DR PROSITE: PS00702; GM_CSF.1.
DR KMW: Polymorphism; Pharmacological.
KM Cytokine; Growth factor; Glycoprotein; Signal; 3D-structure;
FT SIGNAL 1 17
FT CHAIN 18 144
FT FT
FT DISULFID 71 113
FT FT
FT DISULFID 105 138
FT FT
FT CARBOHYD 22 22
FT FT
FT CARBOHYD 24 24
FT FT
FT CARBOHYD 26 26
FT FT
FT CARBOHYD 27 27
FT FT
FT CARBOHYD 44 44
FT FT
FT CARBOHYD 54 54
FT FT
FT VARIANT 115 115
FT FT
FT VARIANT 117 117
FT FT
FT VARIANT 117 117
FT FT
FT TURN 25 27
FT FT
FT HELIX 30 44
FT FT
FT TURN 45 45
FT FT
FT HELIX 50 54
FT FT
FT STRAND 56 60
FT FT
FT HELIX 72 81
FT FT
FT TURN 82 82
FT FT
FT HELIX 85 103
FT FT
FT TURN 104 104
FT FT
FT STRAND 115 119
FT FT
FT HELIX 120 131
FT FT
FT TURN 132 133
SQ SEQUENCE 144 AA; 16295 MW; 75DIE50506BCA7A8 CRC64;

Query Match 18.0%; Score 674; DB 1; Length 144;
Best Local Similarity 99.2%; Pred. No. 9.5e-31;
Matches 127; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 548 AAFARSPSTQPMENHNAIOEARLLNLSRDTAAENETVEYISEMFDLQEPCTQTRL 607
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 17 SARARSPSTQPMENHNAIOEARLLNLSRDTAAENETVEYISEMFDLQEPCTQTRL 76
OY 608 ELYKGLRGLSLTKLKGPLTMASHYKQCHPPTPETSQATQIITFESEKEMDKFLVYIP 667
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 77 ELYKGLRGLSLTKLKGPLTMASHYKQCHPPTPETSQATQIITFESEKEMDKFLVYIP 136
OY 668 DCWEPEQOE 675
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 137 DCWEPEQOE 144

RESULT 11
ERR3_RAT STANDARD: PRT; 1339 AA.
AC 062799; 062955;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)

```

```

DE (c-erbB3).
GN ERB3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
OX NCBI_Taxid-10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Liver;
RX MEDLINE-96096535; PubMed-8522190;
RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;
RT "Cloning of the rat ErbB3 cDNA and characterization of the
RT recombinant protein.";
RL Gene 165:279-284(1995).
RN [2]
RP REVISIONS TO 85; 513 AND 565.
RA Hellyer N.J., Koland J.G.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 922-1097 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-sclatic nerve;
RX MEDLINE-97184212; PubMed-9030624;
RA Carroll S.L., Miller M.L., Frohnert P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuroligins and their putative receptors, Erb2 and
RT Erb3, is induced during Wallerian degeneration.";
J. Neurosci. 17:1642-1659(1997).
CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAK.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U29339; AAC28498.2; -
DR EMBL: U52530; AAC53050.1; -
DR HSSP: P11362; IEGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; FU; 5.
DR SMART: SM00219; TYRKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE-NEG.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 19
FT CHAIN 20 1339
FT DOMAIN 20 643
FT TRANSMEM 644 662
FT DOMAIN 663 1339
FT DOMAIN 183 259
FT DOMAIN 707 964
FT PROTEIN KINASE.

```







RL Genetics 137:531-550(1994).  
 RN [2]  
 RP REVISIONS.  
 RA Clifford R., Schupbach T.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE-85134611; PubMed-2982499;  
 RA Llywne E., Glazer L., Segal D., Schlessinger J., Shilo B.-Z.;  
 RT "The Drosophila EGF receptor gene homolog: conservation of both  
 RT hormone binding and kinase domains.";  
 RL Cell 40:599-607(1985).  
 RN [4]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.  
 RC STRAIN-Oregon-R; TISSUE-Embryo;  
 RX MEDLINE-87002474; PubMed-3093080;  
 RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;  
 RT "Alternative 5' exons and tissue-specific expression of the  
 RT Drosophila EGF receptor homolog transcripts.";  
 RL Cell 46:1091-1101(1986).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I), TISSUE SPECIFICITY, AND MUTATION  
 RP ANALYSIS.  
 RX MEDLINE-99102120; PubMed-9882502;  
 RA Leachkin A.M., Yu S.-Y., Katz J., Baker N.E.;  
 RT "Several levels of EGF receptor signaling during photoreceptor  
 RT specification in wild-type, Ellipse, and null mutant Drosophila.";  
 RL Dev. Biol. 205:129-144(1999).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I).  
 RC STRAIN-Berkley;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaniatis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Landell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abill J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beson K.Y., Bencos P.V., Bernier B.P., Bhandari D., Bolashkov S.,  
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,  
 RA de Paulos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaitai M., Kalish F., Karpen G.H., Ke Z., Kennislo J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,  
 RA Laako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissensbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [7]  
 RP SEQUENCE OF 959-1078 FROM N.A.  
 RC STRAIN-Daekwanryeong;

RX MEDLINE-85137938; PubMed-2983232;  
 RA Wadsworth S.C., Vincent W.S. III, Billoreau-Wentworth D.;  
 RT "A Drosophila genomic sequence with homology to human epidermal  
 RT growth factor receptor.";  
 RL Nature 314:178-180(1985).  
 RN [8]  
 RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION  
 RP ANALYSIS.  
 RX MEDLINE-92038942; PubMed-1936959;  
 RA Raz E., Schejter E.D., Shilo B.Z.;  
 RT "Interallelic complementation among DER/flb alleles: implications for  
 RT the mechanism of signal transduction by receptor-tyrosine kinases.";  
 RL Genetics 129:191-201(1991).  
 RN [9]  
 RP REVIEW.  
 RX MEDLINE-97248481; PubMed-9094709;  
 RA Perrimon N., Perkins L.A.;  
 RT "There must be 50 ways to rule the signal: the case of the Drosophila  
 RT EGF receptor.";  
 RL Cell 89:13-16(1997).  
 CC -1- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,  
 CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAP-  
 CC MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.  
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES, AND FOR THE  
 CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL  
 CC POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE  
 CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOSEROSA  
 CC AND VENTRAL NEUROECTODERMAL CELLS, GERM BAND REPRESSION, CELL FATE  
 CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF  
 CC CUTICLE.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: ISOFORMS TYPE I AND TYPE II ARE TYPE I  
 CC MEMBRANE PROTEINS. ISOFORM TYPE III IS A TYPE III MEMBRANE  
 CC PROTEIN.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; TYPE I (SHOWN HERE), TYPE II AND  
 CC TYPE III. ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: UBICITOUSLY EXPRESSED IN EMBRYOS, IN LARVAE,  
 CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGE OF  
 CC TESTIS AND OVARY, AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST  
 CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW, LEVELS REMAIN HIGH  
 CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN  
 CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX  
 CC AND THORACIC AND ABDOMINAL GANGLIA.  
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF052754; AAC08536.1; -;  
 DR EMBL; AF052753; AAC08536.1; JOINED.  
 DR EMBL; AF052754; AAC08535.1; -;  
 DR EMBL; AF052752; AAC08535.1; JOINED.  
 DR EMBL; K03054; AAAS1462.1; -;  
 DR EMBL; K03417; AAAS1460.1; -;  
 DR EMBL; K03416; AAAS0965.1; -;  
 DR EMBL; K03418; AAAS1461.1; -;  
 DR EMBL; AF109077; AAD26134.1; -;  
 DR EMBL; AF109078; AAD26132.1; -;  
 DR EMBL; AF109082; AAD26132.1; JOINED.  
 DR EMBL; AF109078; AAD26133.1; -;  
 DR EMBL; AF109084; AAD26133.1; JOINED.  
 DR EMBL; AF109079; AAD26130.1; -;  
 DR EMBL; AF109081; AAD26130.1; JOINED.  
 DR EMBL; AF109079; AAD26131.1; -;  
 DR EMBL; AF109083; AAD26131.1; JOINED.  
 DR EMBL; AF109080; AAD26135.1; -;  
 DR EMBL; AF109080; AAD26135.1; -;  
 DR EMBL; AE003454; AAF46732.1; -;

```

DR EMBL: X02293; CAA26157.1; -
DR EMBL: X78920; CAA55523.1; -
DR EMBL: X78918; CAA55521.1; -
DR EMBL: X78919; CAA55522.1; -
DR PIR: A00640; GQFE.
DR HSP: P11362; 1FGK.
DR FlyBase: FBgn0003731; Egfr.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR Pfam: PF00757; Furin-like.1.
DR Pfam: PF01030; Recep_L_domain.2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase.1.
DR SMART: SM00261; Fy; 7.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM.1.
DR Transmembrane; Glycoprotein; Receptor; Phosphorylation; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Signal; Alternative splicing;
KW developmental protein.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 1426 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT DOMAIN 31 868 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 869 889 POTENTIAL.
FT DOMAIN 890 1426 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 938 1198 PROTEIN KINASE.
FT NF_BIND 944 952 ATP (BY SIMILARITY).
FT BINDING 971 971 ATP (BY SIMILARITY).
FT ACT_SITE 1063 1063 BY SIMILARITY.
FT MOD_RES 902 902 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).

Query Match 14.3%; Score 536; DB 1; Length 1426;
Best Local Similarity 36.8%; Pred. No. 3.6e-22;
Matches 112; Conservative 44; Mismatches 118; Indels 30; Gaps 7;

OY 34 ELARGASTVCTGCTDKRLPASPEHLMDRLHYGCGVVOGNLTLPT-NASLSF 92
DB 95 EFVVG---KICIGTSRLSVPSNKEHNYRLRDYRNCYVDGNLTLMPNENLDSF 150
OY 93 LODQEVGVLLAHNVROVPLRLIVGTQF-----EDNALAVLDNGPLNNTTP 147
DB 151 LDNREYTGILLSHVDKVPKLIQINGRLFLSLVEEKALPV----- 198
OY 148 VTGASPGELRELRLSTELIKGVLLQRPOLCYODTILMKDIFHNQDLATLIDTR 207
DB 199 -----TYSKMTLELPDLVDLNGOVGFHNNYNLCMRTIOMSELVSGTDATYYDTAP 254
OY 208 SRACHPSPCKSGRCMGESSEDOQLRTVYCAAGCA--RCKGPLPTDCHEOCAGCTG 265
DB 255 ERECPKHCSTHG-CWGECPKNCQKFSKLTSCPGAGCYGPKPRECHLFCAGCTG 313
OY 266 PKHSDCLAFHFNHSGICELHCPALTYNDTPESMPNPGRTFGSCVYACRYNLTST 325
DB 314 PTQKDCIACNFFDEAVSKCECPMKRYNPTTYVLETNPEGKYAYGATCYKCEP-GHLRL 372
OY 326 DVGS 329
DB 373 DNCA 376

RESULT 15
CSF2_CEREL STANDARD: PRT: 144 AA.
AC P51748:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
DE (Colony-stimulating factor) (CSF).

```

```

GN CSF2.
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9660.
RN
RP SEQUENCE FROM N.A.
RA Lockhart E.A.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U14392; AAA21439.1; -
DR HSP: P04141; 2GMF.
DR InterPro: IPR000773; GM_CSF.
DR Pfam: PF01109; GM_CSF.1.
DR PRINTS: PR00693; GMCSFACITOR.
DR PRODOM: PD007349; GM_CSF.1.
DR SMART: SM00040; CSF2.1.
DR PROSITE: PS00702; GM_CSF.1.
KW Cytokine; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 17 BY SIMILARITY.
FT CHAIN 18 144 GRANULOCYTE-MACROPHAGE COLONY-STIMULATING
FT FT FACTOR.
FT DISULFID 71 113 BY SIMILARITY.
FT DISULFID 105 138 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 44 44 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 144 AA; 16283 MW; 1F5FF5FD03C94394 CRC64;

Query Match 14.1%; Score 529; DB 1; Length 144;
Best Local Similarity 76.6%; Pred. No. 7.9e-23;
Matches 98; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 548 AAPRSPSPSTQPEWHYNAIOEARRLNLSRDTAAENKETEVEVISEMFDOEPTCLOTRL 607
DB 17 SAPRSPSPPTTRPQONHDAIKELSLNHSDDTAAYVNEVEVSEMFDOEPTCLOTRL 76
OY 608 ELYKQGLRGLTKLGLTMMASHYKQCPPTPETSCATQIITFESKRNLDKFLVYFP 667
DB 77 KLYKQGLRGLSTLSGLTMMARHYEQHPDETSCETQITFKSPKRNLDKFLFIPE 136
OY 668 DCWEPPGE 675
DB 137 DCWEPPAK 144

Search completed: April 28, 2003, 13:40:04
Job time : 21.4231 secs

```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:38:09 ; Search time 25.3846 seconds

(without alignments)  
2613.108 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739  
Sequence: 1 MRAPLILANASLSGLF.....EPVQGNAPPPAAHHHHH 690

Scoring table: BIOSIM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2405.5	64.3	1255	1 A24571	protein-tyrosine k
2	1950.5	52.2	1260	1 TYRNU	protein-tyrosine k
3	1935	51.8	1254	2 I48161	p-185 precursor
4	755.5	20.2	1223	1 TVCHLV	epidermal growth f
5	754.5	20.2	527	2 A42032	epidermal growth f
6	721	19.3	1210	1 GQHUE	epidermal growth f
7	721	19.3	1308	2 A47253	epidermal growth f
8	717.5	19.2	1210	2 A53183	epidermal growth f
9	714.5	19.1	644	2 A56325	epidermal growth f
10	697	18.6	1342	2 A56223	kinase-related tra
11	674	18.0	144	1 FQHUCM	granulocyte-macrop
12	662	17.7	1339	2 JCA387	epidermal growth f
13	655.5	17.5	1166	1 S06142	protein-tyrosine k
14	549	14.7	144	2 JH0469	granulocyte-macrop
15	545	14.6	144	1 A61632	granulocyte-macrop
16	536	14.3	843	2 A27131	epidermal growth f
17	481.5	12.9	143	1 FQBOGM	granulocyte-macrop
18	476.5	12.7	144	2 A44936	granulocyte-macrop
19	441	11.8	127	2 I46269	granulocyte-macrop
20	404.5	10.8	1323	2 E88257	protein let-23 (lm
21	404.5	10.8	1374	2 S70712	protein-tyrosine k
22	390.5	10.4	1369	2 S70713	protein-tyrosine k
23	375.5	10.0	1330	1 GQFPE	epidermal growth f
24	374	9.1	153	1 FQMSGM	granulocyte-macrop
25	341	9.1	366	2 D45558	epidermal growth f
26	341	8.9	1717	2 A45558	epidermal growth f
27	331	8.9	333	2 B45558	epidermal growth f
28	331	8.9	342	2 C45558	epidermal growth f
29	317.5	8.5	1363	2 T43220	insulin-like growt

30	273.5	7.3	1477	2 T18534	protein-tyrosine k
31	270.5	7.2	2101	2 S57245	insulin receptor (
32	270.5	7.2	2148	1 A56081	insulin receptor
33	256.5	6.9	1382	1 INHUR	insulin receptor p
34	252.5	6.8	1372	2 A34157	insulin receptor p
35	252	6.7	1607	2 T43212	insulin-like growt
36	249	6.7	1383	2 A36080	insulin receptor p
37	246.5	6.6	1300	2 A36502	insulin receptor p
38	246	6.6	1367	1 IGHUR1	insulin-like growt
39	238	6.4	340	2 B47417	insulin receptor-r
40	235.5	6.3	183	2 JH0803	tyrosine kinase re
41	235.5	6.3	1371	2 A33837	insulin-like growt
42	231.5	6.2	1268	2 B36502	insulin receptor-r
43	223.5	6.0	329	2 A48805	insulin-like growt
44	207	5.5	1390	2 T30346	insulin receptor h
45	188	5.0	1846	2 T42047	insulin receptor h

#### ALIGNMENTS

##### RESULT 1

A24571 protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human

N:Alternate names: C-erb-B-2 protein precursor; Kinase-related transforming protein e

C:Species: Homo sapiens (man)

C:Date: 25-Oct-1987 #sequence\_revision 06-Dec-1996 #text\_change 11-Jun-1999

C:Accession: A24571; A23491; A44188; B44188; I55509; I57622

R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Salto, T

Nature 319, 230-234, 1986

A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growl

A:Reference number: A24571; MUID:66118663; PMID:3003577

A:Accession: A24571

A:Molecule type: mRNA

A:Residues: 1-1235 <YAM>

A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198

R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.

Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985

A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid

A:Reference number: A25491; MUID:86016729; PMID:2955967

A:Accession: A25491

A:Molecule type: DNA

A:Residues: 737-1031 <SEM>

A:Cross-references: GB:M1787; NID:g182163; PIDN:AAA35808.1; PID:g553282

A:Accession: B44188

A:Molecule type: mRNA

A:Residues: 1-517, 'RALI', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>

A:Cross-references: GB:M1730; NID:g183986

R:King, C.R.; Kraus, M.H.; Aaronson, S.A.

Science 229, 974-976, 1985

A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.

A:Reference number: I59509; MUID:85272597; PMID:2992089

A:Accession: I59509

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 837-909 <REX>

A:Cross-references: GB:I29395; NID:g459807; PIDN:AAA35809.1; PID:g459808

R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.

Mol. Cell. Biol. 7, 2597-2601, 1987

A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio

A:Reference number: I57622; MUID:87286898; PMID:3039351

A:Accession: I57622

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-191 <TAI>

A:Cross-references: GB:M16792; NID:9183983; PIDN:AAA8637.1; PID:9553332  
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
 C:Genetics:  
 A:Gene: GDB:ERBB2; NGL; NEU; HER-2  
 A:Cross-references: GDB:120613; OMIM:164870  
 A:Map position: 17q21.1-17q21.1  
 A:Introns: 25/1; 75/3; 147/1; 883/3  
 A:Note: the list of introns is incomplete  
 C:Function:  
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-125/Domain: protein-tyrosine kinase erbB2 #status predicted <MAT>  
 F:22-653/Domain: extracellular #status predicted <EXT>  
 F:70-304/Domain: EGF receptor extracellular domain repeat <EB1>  
 F:395-605/Domain: EGF receptor extracellular domain repeat <EB2>  
 F:654-675/Domain: transmembrane #status predicted <TM>  
 F:676-1253/Domain: intracellular #status predicted <INT>  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:126-734/Region: protein kinase ATP-binding motif  
 F:68-124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:753/Active site: lys #status predicted  
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 64.3%; Score 2405.5; DB 1; Length 1255;  
 Best Local Similarity 41.1%; Pred. No. 9,7e-126;  
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGAASVCTGTMKRLRPASETHLMDLRHLYGCGVYVGNLELYLPTNASLSFLQ 94  
 DB 16 LPPGAASTQVCTGDMKRLRPASETHLMDLRHLYGCGVYVGNLELYLPTNASLSFLQ 75  
 QY 95 DIOEVGVYLIANOVQVQLRLRYRGQLPEDNALVLDNGDPLNNTPTTGASPG 154  
 DB 76 DIOEVGVYLIANOVQVQLRLRYRGQLPEDNALVLDNGDPLNNTPTTGASPG 135  
 QY 155 GLRELQRLSLLEILKGGVLIQRNQLCYODTILMKDIFRHNQALTLIDITNRSRACHP 214  
 DB 136 GLRELQRLSLLEILKGGVLIQRNQLCYODTILMKDIFRHNQALTLIDITNRSRACHP 195  
 QY 215 SPKCKGRGCGESSEDCQSLTRYACAGGACRCKPRLTDCCHQCAAGCTGPKHSDCLAC 274  
 DB 196 SPKCKGRGCGESSEDCQSLTRYACAGGACRCKPRLTDCCHQCAAGCTGPKHSDCLAC 255  
 QY 275 LHFHSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYLYLSTDVSGSLVC 329  
 DB 256 LHFHSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYLYLSTDVSGSLVC 315  
 QY 330 ----- 339  
 DB 316 PLNHQEVTAADGTQRCCKSPCARVCYGLMEHLREVRVAVTSANIOEFAKCKKIFGSLA 375  
 QY 330 ----- 329  
 DB 376 FLPESEFDGDAANTAPLQPEOLQVFTLEETIGLYLISAMPDSLPLDSVFNQNLQVIRGI 435  
 QY 330 ----- 329  
 DB 436 LHHGANSVLTLQGLGISWGLRLSRELGLALIHNTHTLFCVHTVPWDQLFRNPQALLH 495  
 QY 330 ----- 329  
 DB 496 TANRPDECVGEGILACHQCLARGHCWGPPTQVCVNCQFLRGOECVEECRVLDGLPREYV 555  
 QY 330 ----- 329  
 DB 556 NARHCLPCHPECOPIQNSVTCFGEPAQCVAACHYKDPFCVACRPSGVKPLSLYPTWK 615  
 QY 330 ----- 329

DB 616 FPBEGACQCPINCHSCVDLDDKCGPAQRASPLTISAVVGLLVVLCVGEGLI 675  
 QY 330 ----- 329  
 DB 676 KRQOKIRKXTMRLLQETELVEPLTPSGAMPQAOIRILKETELRKVKVLSGAGTGY 735  
 QY 330 ----- 329  
 DB 736 KGIWIPDENVKIPVAIKVLRNTSPKANKEILDEAVVMAGVSPYVRLGLTSTVQ 795  
 QY 330 ----- 329  
 DB 796 LVTQLMPYGLLDHVRNKRGLSODLLNMCQIANGMSYLEDRVLRDARNAVLYKS 855  
 QY 330 ----- 329  
 DB 856 PNHKIDPGLRLDIDETEHADGKVPDKMALESILRRFTHQSDVKVGVYVWEL 915  
 QY 330 ----- 329  
 DB 916 MTEGAKPYDGIPIAREIPDLLEKGERLPPPICTIDVYIMVKWMIDSECRPRRELVE 975  
 QY 330 ----- 329  
 DB 976 FSRMARDQRFVVIQNEDELGPASPLDSTFYRSLLEDDMDGLVDAEYLVPOGCFPCPP 1035  
 QY 330 ----- 329  
 DB 1036 APGAGVGHRRHSSSTRSGGDLTLGLEPSEEARSPAPSEAGSDVFDGDLGMA 1095  
 QY 388 KGLQSLPTHDSPLOXYSEDPVPLPSETDGYAAPLTCSPQPEYVNOVDVPPSPREG 447  
 DB 1096 KGLQSLPTHDSPLOXYSEDPVPLPSETDGYAAPLTCSPQPEYVNOVDVPPSPREG 1155  
 QY 448 PLPAARAGATLERAKTLRSGKGVYKDVAFGAVENPEYLPPOGGAAPPPAPFSP 507  
 DB 1156 PLPAARAGATLERAKTLRSGKGVYKDVAFGAVENPEYLPPOGGAAPPPAPFSP 1215  
 QY 508 AFNLLYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 546  
 DB 1216 AFNLLYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 1254

## RESULT 2

TVRTNU  
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 11-Jun-1999  
 C:Accession: A24562; A61204  
 R:Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.  
 Nature 319, 226-230, 1986  
 A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein  
 A:Reference number: A24562; MUID:86118662; PMID:3945311  
 A:Accession: A24562  
 A:Molecule type: mRNA  
 A:Residues: 1-1260 <BAR>  
 A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746  
 R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohe  
 Carcinogenesis 12, 1975-1978, 1991  
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals n  
 2-thiazolyl]formamide or N-methyl-N-nitrosourea.  
 A:Reference number: A61204; MUID:92035293; PMID:1682063  
 A:Accession: A61204  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 637-663, 'V', 665-702 <MAS>  
 A:Note: authors translated the codon GCA for residue 25 as Val  
 C:Genetics:  
 A:Gene: neu  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

F:658-680/Domain: transmembrane #status predicted <TM>  
F:723-988/Domain: protein kinase homology <KIN>  
F:731-739/Region: protein kinase ATP-binding motif  
F:771-191,263,555,576,634/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:691/Binding site: phosphate (Tmr) (covalent) #status predicted  
F:759/Active site: Lys #status predicted  
F:882,1277,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match	52.28;	Score 1950.5;	DB 1;	Length 1260;
Best Local Similarity	34.18;	Pred. No. 1.6e-100;		
Matches 433;	Conservative 24;	Mismatches 72;	Indels 741;	Gaps 3

Qy	6	LLLRAASTLSIGFLEFLEFMDRSLVAKELRGAASTOVGTGDMKLRLPASPETHDML	65
Dp	2	IMELMAACRBMGFLAL-----LPRGIAAGVOVCTGDMKLRLPASPETHDML	49
Qy	66	RHLVYOGCOVVOGNILETLYLPTPNASLSFLODIOEVGYVLAHNOVOVPLRLRIVRGTO	1235
Dp	50	RHLVYOGCOVVOGNILETLYVPPANASLSFLODIOEVGYMLAHNOVKRVPRLRIVRGTO	1090
Qy	126	LFEDNYALAVLDNDPDLNNTPTV- GASPGLRLEDOBSLTELKGVLOLRNOQLCYOD	1840
Dp	110	LFEDNYALAVLDNDPDLNNTPTV- GASPGLRLEDOBSLTELKGVLOLRNOQLCYOD	1690
Qy	185	TLIMKDFHKNNQOLALTIDITNRSRACHPCSPMKGSGSCMESSSEDCOSILRTVCAGCA	2440
Dp	170	MYLMKDWFRKNNQOLAPVIDIDITNRSRACHPCSPACKDNHCHMESPEDOILIGTICTSGCA	2220
Qy	245	RCKGRPLPTDCCHECOCAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMPNP	3040
Dp	230	RCKGRPLPTDCCHECOCAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMNP	2890
Qy	305	EGRTFGASCTVACPRNYLSTDVDS-----	3220
Dp	290	EGRTFGASCTVACPRNYLSTDVDS-----	3490
Qy	330	-----	3220
Dp	350	GMEHLRGARAITSDNVQEFDCCKKIFGSLAFLPESFDGDPSSGIAPLRPEOLOFETLEE	4090
Qy	330	-----	3220
Dp	410	ITGYLYISAMPDLSRLDLSVFQNLRIINGRIHLHDGAYSLLTQGLGISHLSLRELSGL	4460
Qy	330	-----	3220
Dp	470	ALIHNRNAHLCFVHTVPMDOLEFRNPHQALLHSGNRPREDLCVSGSLVCNSLCAHGCWPG	5220
Qy	330	-----	3220
Dp	530	PTOCVNCSHFLRGQCEVCECHVKMGKLPREYVSDKRCLEPCHECOPONSSSETFCSEADOC	5890
Qy	330	-----	3220
Dp	590	AACAHYKDSGSSGVARCPBGVAKPDLSTYMPIMKYPRBEGICQPCPLNCTHNSCVDLDERCPA	6440
Qy	330	-----	3220
Dp	650	EORASPVTFIATVEGVLLFLVWVVGILIKRRRQKIRKTYMRLLQETELVEPLPSC	7090
Qy	330	-----	3220
Dp	710	AMPNOAMRILKETELRKVKVLSGAFGTYYKGIWIDGENVKIPVAIKYLRENTSPPAN	7650
Qy	330	-----	3220
Dp	770	KEILDEALVMAAGVSPYVSRLLGICLSTVOVLVQMLPGLCLDHVREHNRGLSGDLLN	8220
Qy	330	-----	3220
Dp	830	MCVOJAKGNSLYLEDEVRLVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEVHADGKV	8890
Qy	330	-----	3220

Db	890	PIKMALESILRRRFTHQSDVMSGYVWMLMTGAKPDIQIPAREIPDLLEKGFRLP	949
QY	330	-----	329
Db	950	PICTIDVYIMVWKCMIDSECRPRFRELYSEFSHMARDQRFVYIQNEDLSPSFSMDSTF	1009
QY	330	-----GAGCVHHRHSSSTRSGGDLTGLE	356
Db	1010	YRSLLEDDMDGLDAEEYLVPGQGFSPDPTPETGSTAHNRHSSSTRSGGDLTGLE	1069
QY	357	PSEEARPRCLAPSEAGSDVFDGDLGMAKGLQSLRPHDSPLRQSDPYPLPSET	416
Db	1070	PSEEGPRSPAPSEAGSDVFDGDLAMGYKGLQSLSPHDSLQKXSDPTPLRPET	1129
QY	417	DGYVAPLTCSPOPEYVNOPDVRPQPSBREGPLPAARPAGATTLERAATLSFGKGVKDV	476
Db	1130	DGYVAPLACSPQPEYVNOSEVOQPRLPTEGRLPPVRPAGATTLERPKTLSPGKGVKDV	1189
QY	477	FAFGAVENPEYLTLPQGGAAFPQHPRAFSFAFNLYYWDODPPERGAPSTFNGTPTAE	536
Db	1190	FAFGAVENPEYLTVPREGTASPRHPSPAFSPFNLYYWDONSSGQPPSPNFGTPTAE	1249
QY	537	NPEYLGLDVP	546
Db	1250	NPEYLGLDVP	1259

```

RESULT 3
148161
P:185 precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
C:Accession: I48161
R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishika
Gene 140, 251-255, 1994
A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
A:Reference number: I48161; MUID:94193007; PMID:7908275
A:Accession: I48161
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1254 <RES>
A:Cross-references: GB:D16295; NID:g493236; PIDN:BA003801.1; PID:g747595
C:Genetics:
A:Gene: neu
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP
E:718-963/Domain: protein kinase homology <KIN>
E:776-794/Region: protein kinase ATP-binding motif

```

	Query Match	Best Local Similarity	Matches	Conservative	Score 1935	Pred. NO. 1.le-99	Length 1254	DB 2				
			423	32	Mismatches 66	Indels 728	Gaps 2					
QY	25	WLDKRVLAKE	LARGAAS	TVCTGTCTDMKRL	LRPASPENHIDML	KRLHYGGCGQVQGNLELTYL	84					
Db	6	WCGMGLLALL	SPGASG	VGCTGTDMKRL	LRPASPENHIDI	VHRLVGGCGVQGNLELTYL	65					
QY	85	PTNASTLSE	LQDIOEQV	GYVLIAHNO	VRQPLRLR	IVRGTCTFEDNYALAVL	DNGDPLNN 144					
Db	66	PANATLST	LQDIOEQV	GYVLI	IAHNSQVR	HAHPLQRLR	IVRGTCTFEDKXALAVL	DNRDPLDN 125				
QY	145	TTPTVYGAS	PGGILREQL	SLSTLEIL	KGVLQ	RNPQLCTYD	TLIMKDI	EFKKNQALATLTD 204				
Db	126	VTTAAGRPE	GRPEGRLE	QTLRSL	TEILK	GGVLI	IRKNPQLCTYD	TALMKVFERKKNQALAVDID 185				
QY	205	TNRSRAC	HPSCB	MCKSGRS	GESSESD	CSLITRTV	CAGCARGC	KGPLPTPCCH	DEQCAAGT 264			
Db	186	TNRSRAC	PPCAPAC	CDNMCMGAS	PRD	CDQTLTG	TIA	PAVAPARARL	PTDC	CHDCAAGT 245		
QY	265	GPKHSD	CLACL	AFHNSG	ICELC	HPALVY	NTD	FESMP	PRBGRYTF	GASCV	PAAPNYLS 324	
Db	246	GPKHSD	CLACL	AFHNSG	ICELC	HPALVY	NTD	FESMP	PRBGRYTF	GASCV	TPCPNYLS 305	



C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 18-Jun-1999  
 C:Accession: A42032  
 R:Flickinger, T.W.; Mahle, N.J.; Kung, H.J.  
 Mol. Cell Biol. 12, 883-893, 1992  
 A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, t  
 A:Reference number: A42032; MUID:92123214; PMID:1732751  
 A:Accession: A42032  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-527 <Full>  
 A:Cross-references: GB:M7637; NID:9211737; PIDN:AAA48759.1; PID:9211738  
 A:Experimental source: liver  
 A:Note: sequence extracted from NCBI backbone (NCBI:76892, NCBI:76893)  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; growth factor receptor

Query Match 20.2%; Score 754.5; DB 2; Length 527;  
 Best Local Similarity 42.4%; Pred. No. 1.1e-34;  
 Matches 160; Conservative 56; Mismatches 126; Indels 35; Gaps 10;

Oy 19 LELFPMIDRSVLAKELARGASTGVCTGDMKRLRLPASPEHLDMLRHLVYGCQVQGN 78  
 18 LVLLLLLRVALCS-----ANEKRVCCGTNNKLTQLGHVEHFSLOMNNCEVVLN 73  
 Oy 79 LELTLPNASTLFQIDQEVGYVLIANQVQVPLQRLRIIVRGTOLEFEDNYALAVLDN 138  
 74 LEITVEHNRDLTFLKLTQEVAGVYLIANMVDYPLEMLQIIRGNVLYDNEFALAVLSN 133  
 Oy 139 GDPNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDIFKNNOL 198  
 134 YH-MNKTO-----GLRELPMKRLSELINGVAKISNNPKLCNMDVYLMNDIIDSFRK- 183  
 Oy 199 ALTLID-TNRSRACHPCSPMCKSGRSCSSSDCSLTPTVAGCA-RCCKPLPTDCH 256  
 194 PLTVLDPSNLSCKPCPCNCTEDHCWGAGEQNCQTLTKVICAQCSGCKRCKVSDCH 243  
 Oy 257 EDCAGACTGPKHSDCLACLFHNSICELHPCALVTYNTDFEESNPFGRTFCASCVT 316  
 244 NCGAAGCTGPRSDCLACKFRDQATKDTCPPLVLYNTTYQMVDNPEGKTSFQATCR 303  
 Oy 317 ACPYVLTLDVSGGAGVNHHRSSSTRSGGDLTLGLEPSEEPAPSLAPSEAGSD 376  
 304 ECPHYVVTDHGSCV-----RSCNMDT-----YEVEENGVRK-CKKCDGLCSK 345  
 Oy 377 VFDGDLGMAKGLSL 393  
 346 VCNQ-IGIGELKGLSL 361  
 Db

RESULT 6  
 GOHUE  
 epidermal growth factor receptor precursor - human  
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB  
 C:Species: Homo sapiens (man)  
 C:Date: 15-Nov-1984 #sequence\_revision 27-Nov-1985 #text\_change 11-Jun-1999  
 C:Accession: A00641; A25772; S30024; A38672; A33615; A23062; A05281; A60143; A33  
 R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y  
 rg, P.H.  
 Nature 309, 418-425, 1984  
 A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of  
 A:Reference number: A00641; MUID:84219729; PMID:6328312  
 A:Accession: A00641  
 A:Molecule type: mRNA  
 A:Residues: 1-1210 <Full>  
 A:Cross-references: EMBL:X00586; NID:q31113; PIDN:CAA25240.1; PID:q757924  
 A:Note: the authors translated the codon AAG for residue 540 as Asn  
 R:Shih, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985  
 A:Title: Characterization and sequence of the promoter region of the human epidermal gro  
 A:Reference number: A25772; MUID:85270438; PMID:2991899  
 A:Accession: A25772  
 A:Status: translation not shown  
 A:Molecule type: DNA

A:Residues: 1-29 <ISH>  
 A:Cross-references: GB:M11234; NID:q181981; PIDN:AAA52370.1; PID:q553272  
 R:Haley, J.; Whittle, N.; Bennett, P.; Kirschington, D.; Ulrich, A.; Waterfield, M.  
 Oncogene Res. 1, 375-396, 1987  
 A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification  
 A:Reference number: S30024; MUID:88217333; PMID:3329716  
 A:Accession: S30024  
 A:Molecule type: DNA  
 A:Residues: 1-29 <HA2>  
 A:Cross-references: EMBL:X06370; NID:q31118; PIDN:CAA29668.1; PID:q31119  
 R:Haley, J.D.; Waterfield, M.D.  
 J. Biol. Chem. 266, 1746-1753, 1991  
 A:Title: Constitutive effects of de Novo transcription and premature transcript termi  
 A:Reference number: A38672; MUID:91107677; PMID:1988448  
 A:Accession: A38672  
 A:Molecule type: DNA  
 A:Residues: 1-29 <HAL>  
 A:Cross-references: GB:M38425; NID:q181977; PIDN:AAA63171.1; PID:q553271  
 A:Experimental source: carcinoma cell line A431-7  
 R:Xu, Y.; Shih, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;  
 Nature 309, 806-810, 1984  
 A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN  
 A:Reference number: A00642; MUID:84245835; PMID:6330563  
 A:Accession: A00642  
 A:Molecule type: mRNA  
 A:Residues: 'RCAMRRA', 150-187, 'KSVIOAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-  
 '798-799, 'ND', 802-811, 'R', 813-942 <XUV>  
 A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF re  
 R:Lin, C.R.; Chen, W.S.; Krulger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma,  
 Science 224, 843-848, 1984  
 A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplificacl  
 A:Reference number: A43615; MUID:84196372; PMID:6326261  
 A:Accession: A43615  
 A:Molecule type: mRNA  
 A:Residues: 713-964 <LIN>  
 A:Experimental source: epidermoid carcinoma cell line A431  
 R:Slimem, F.A.; Gope, M.L.; Schultz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.  
 Biochem. Biophys. Res. Commun. 124, 123-132, 1984  
 A:Reference number: A23062; MUID:85046483; PMID:6093780  
 A:Accession: A23062  
 A:Molecule type: mRNA  
 A:Residues: 1028-1210 <SIM>  
 R:Weber, W.; Gull, G.N.; Speiss, J.  
 Science 224, 294-297, 1984  
 A:Reference number: A05281; MUID:84172183; PMID:6324343  
 A:Accession: A05281  
 A:Molecule type: protein  
 A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>  
 R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.  
 J. Biol. Chem. 260, 5205-5208, 1985  
 A:Title: Identification of residues in the nucleotide binding site of the epidermal g  
 A:Reference number: A60143; MUID:85182650; PMID:2985580  
 A:Accession: A60143  
 A:Molecule type: protein  
 A:Residues: 740-744, 'X', 746-747 <RUS>  
 R:Procakowski, B.; Mosig, G.; Cohen, S.  
 Nature 309, 270-273, 1984  
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and supe  
 A:Reference number: A38023; MUID:84191554; PMID:6329948  
 A:Contents: annotation; receptor activity  
 A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA  
 R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C  
 Cell 59, 33-43, 1989  
 A:Title: Functional independence of the epidermal growth factor receptor from a domai  
 A:Reference number: A3331; MUID:90003233; PMID:2790960  
 A:Contents: annotation; internalization signal  
 C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-recepto  
 C:Genetics:  
 A:Gene: GDB:EGFR  
 A:Cross-references: GDB:120610; OMIM:131550  
 A:Map position: 7p12.3-7p12.1  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos





A:Cross-references: GB:L068664; NID:g193001; PIDN:AAA53029.1; PID:g567201  
 C:Genetics:  
 A:Gene: EGFR  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphop  
 F:1-24/Domain: signal sequence #status predicted <Stic>  
 F:648-670/Domain: transmembrane #status predicted <TM>  
 F:712-977/Domain: protein kinase homology <KIN>  
 F:720-728/Region: protein kinase ATP-binding motif  
 F:680-695/Binding site: phosphate (Thr) (covalent) #status experimental  
 F:697-1070/Binding site: phosphate (Ser) (covalent) #status experimental  
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental  
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental  
 F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 19.2%; Score 717.5; DB 2; Length 1210;  
 Best Local Similarity 46.8%; Pred. No. 3.1e-32;  
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;

QY 36 ARGAA--STOVCTGTDMLRLPASPEHLDMLRHLYOGCOVVOGNLELTYLPTNASLFL 93  
 DB 20 AAGALEEKVVCQGTNRLLTQLTFEDHFLSLDRMYNCEVVLGNLETTYQKNYDSEL 79  
 QY 94 QDIOEVGYVLIANOVROYPLQRLIRVGTQDPEFNALAVLNDGDPNLNTPVTCASP 153  
 DB 80 KTIQEVAGYVLIANTVERIPLENLQIIRGNALYENTYALALISN-----YGTNR 129  
 QY 154 GGLREQLRLSTELKGVLIQIRNPOLCYODTILMKDI----FHKNNQLALTLIDTRRSR 209  
 DB 130 TGIREFLPMRLQILIGAVFSSNNPLICNMDTQMRITQVNMFSNMSMDL---OSHP 185  
 QY 210 ACHPCSPMKSGRSGWSESSDCSLRTVYAGGCA--RCKGPLPTDCHEQCAAGCTGPKH 268  
 DB 186 SCRCPCDSCNGSCWGGEENCKLTKIICAOQCSHRCRRSPSDCHNCNCAAGCTGPRE 245  
 QY 269 SDCGLACHFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCYTACPYNLSTDVG 328  
 DB 246 SDCLVCKCFQDEATCKDKTCLPLMLYNPTTYQMDVNPFGKYSFGATCVKCPRNYYVDHG 305  
 QY 329 S 329  
 DB 306 S 306

RESULT 9  
 A36325  
 epidermal growth factor receptor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 25-Jan-1991 #sequence\_revision 25-Jan-1991 #text\_change 10-Oct-1997  
 C:Accession: A36325  
 R:Fetch, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Earp, H.S.  
 Mol. Cell. Biol. 10, 2973-2987, 1990  
 A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded b  
 A:Reference number: A36325; MUID:90258888; PMID:2342466  
 A:Accession: A36325  
 A:Status: preliminary  
 A:Superfamily: ATP; phosphotransferase  
 A:Keywords: ATP; phosphotransferase  
 A:Residues: 1-644 <DET>  
 A:Cross-references: GB:M37394  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 19.1%; Score 714.5; DB 2; Length 644;  
 Best Local Similarity 37.2%; Pred. No. 2.3e-32;  
 Matches 169; Conservative 60; Mismatches 158; Indels 67; Gaps 13;

QY 30 VLAKELARGAA--STOVCTGTDMLRLPASPEHLDMLRHLYOGCOVVOGNLELTYLPTN 87  
 DB 14 LLAALCAAGALBEKKVCGCTSNRLQLGFEDHFLSLDRMYNCEVVLGNLETTYQKN 73  
 QY 88 ASLSFLQDIOEVGYVLIANOVROYPLQRLIRVGTQDPEFNALAVLNDGDPNLNTP 147  
 DB 74 YDLSFLKTIQEVAGYVLIANTVERIPLENLQIIRGNALYENTYALALISN----- 124

QY 148 VTGASPGRELQRLSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQLALTLIDTNR 207  
 DB 125 -YGTNKTGRELRLMRNLQELIGAVFSSNNPLICNMDTQMRITV-QDVLSSMSMDYOR 182  
 QY 208 S-RACHPCSPMKSGRSGWSESSDCSLRTVYAGGCA--RCKGPLPTDCHEQCAAGCTG 265  
 DB 183 HLNGCPKCDSPCNGSCWGGEENCKLTKIICAOQCSHRCRRSPSDCHNCNCAAGCTG 242  
 QY 266 PKISDCLACHFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCYTACPYNLST 325  
 DB 243 PRSDCLVCHRFQDEATCKDKTCLPLMLYNPTTYQMDVNPFGKYSFGATCVKCPRNYYVT 302  
 QY 326 DVCSGAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPLAPSGCAGSDVFDGDLGMC 385  
 DB 303 DHG-----SCVACCPDIY-----EVEEDGVSKCKKCDKPCAKVNG-IGIG 343  
 QY 386 AAKGLQSLPTHPDPSPLQRYSEDPVPLPSETDGVVAPLTCSPQEPVYNQDPVRPQPSPR 445  
 DB 344 EFRDITLST---NMNIRKHFYCTAI---SGDLHLPLVA-----FKGDSFTTRPPLDPR 390  
 QY 446 EGPPLPAAPAGATLERAKTILSPGKNGVYKDYFAF 479  
 DB 391 E-----LEILKT-----VKEITGF 404

RESULT 10  
 A36223  
 kinase-related transforming protein (erbB3) (EC 2.7.1.-) precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Oct-1991 #sequence\_revision 13-Jan-1993 #text\_change 17-Nov-2000  
 C:Accession: A36223; 159164  
 R:Kraus, M.H.; Issing, W.; Miki, T.; Popescu, N.C.; Aaronson, S.A.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989  
 A:Title: Isolation and characterization of ERBB3, a third member of the ERBB/epiderma  
 A:Reference number: A36223; MUID:90083234; PMID:2687875  
 A:Accession: A36223  
 A:Status: preliminary  
 A:Superfamily: ATP; phosphotransferase  
 A:Keywords: ATP; phosphotransferase  
 A:Residues: 1-1342 <KRA>  
 A:Molecule type: mRNA  
 A:Cross-references: GB:M29366  
 R:Plowman, G.D.; Whitney, G.S.; Neubauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro,  
 Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990  
 A:Title: Molecular cloning and expression of another epidermal growth factor receptor  
 A:Reference number: 159164; MUID:90311312; PMID:2164210  
 A:Accession: 159164  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Superfamily: ATP; phosphotransferase  
 A:Keywords: ATP; phosphotransferase  
 A:Residues: 1-559, 'G', '561-957', 'F', '959-1063', 'G', '1065-1342 <RES>  
 A:Cross-references: GB:M34309; NID:g183990; PIDN:AAA35979.1; PID:g106841  
 C:Genetics:  
 A:Gene: GDB:ERBB3; HER3  
 A:Cross-references: GDB:119880; OMIM:190151  
 A:Map position: 12q13-12q13  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
 C:Keywords: ATP; phosphotransferase  
 F:707-972/Domain: protein kinase homology <KIN>  
 F:715-723/Region: protein kinase ATP-binding motif

Query Match 18.6%; Score 697; DB 2; Length 1342;  
 Best Local Similarity 44.6%; Pred. No. 4.8e-21;  
 Matches 139; Conservative 45; Mismatches 98; Indels 30; Gaps 7;

QY 16 LGFLFLFFFLDLSVLAKELARGAA--STOVCTGTDMLRLPASPEHLDMLRHLYOGC 72  
 DB 10 LGFLF-----SLARGSEVGNSSQAVCPGTUNGSLVGDADNVOYQTLTKLTERC 56  
 QY 73 QVVOGNLELTYLPTNASLFLQDIOEVGYVLIANOVROYPLQRLIRVGTQDPEFNVA 132  
 DB 57 EYVWGNLEIYLTGNMADSLQWIREVTGYVLAMNFSFLPLPNLRAVVGQVYDCKFA 116  
 QY 133 LAVLNDGDPNLNTPPYTCASPGGLRELQRLSLTEILKGVLIQIRNPOLCYODTILMKDIF 192  
 DB 133 LAYLNDGDPNLNTPPYTCASPGGLRELQRLSLTEILKGVLIQIRNPOLCYODTILMKDIF 192

Db 117 IFPM-----LNYNT-----NSSHALRLQLTQLTEILISGVIYEKNDLCHMDITDMRDIV 167  
 Qy 193 HKNNQALATLIDTRNSRACHPCSPMCKSGRSGESSEDSCSLTRTYCAGGC-ARCKGPLP 251  
 Db 168 RDRD---AEIVVDNGNSCPCHVEYCKG-RCKWGSSEDCTLRKTKCAPOCNCHGCEFPNP 223  
 Qy 252 TDCCHDCAAGCTGPKRISDCLACLFHNSGICELHCPALVTYNTDFESMPNPGRYTGC 311  
 Db 224 NQCHCEACAGCGSPDPTDCEACRHFNDSGACVPCRPQPLVYNKLFQLEPNPHTKYQYG 283  
 Qy 312 ASCVTACPYNYL 323  
 Db 284 GVCVASCPHFV 295

RESULT 11  
 FOHUGM  
 granulocyte-macrophage colony-stimulating factor precursor (validated) - human  
 N:Alternate names: colony-stimulating factor 2; GM-CSF  
 C:Species: Homo sapiens (man)  
 C>Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 08-Dec-2000  
 C:Accession: C24636; 159065; A25169; A01853; A44175; JCI090  
 R:Myatake, S.; Otsuka, T.; Yokota, T.; Lee, F.; Aral, K.  
 EMBL J. 4. 2561-2568, 1985  
 A>Title: Structure of the chromosomal gene for granulocyte-macrophage colony stimulating  
 A:Reference number: A91015; MUID:86030234; PMID:3876930  
 A:Accession: C24636  
 A:Molecule type: DNA  
 A:Residues: 1-144 <MIV>  
 A:Cross-references: EMBL:X03021; NID:g931858; PIDN:CAA26822.1; PID:g931859  
 R:Kaushtansky, K.; O'Hara, P.J.; Berkner, K.; Segal, G.M.; Hagen, F.S.; Adamson, J.W.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 3101-3105, 1986  
 A>Title: Genomic cloning, characterization, and multilineage growth-promoting activity  
 A:Reference number: 159065; MUID:86205844; PMID:3486413  
 A:Accession: 159065  
 A>Status: translated from GB/EMBL/DBDUT  
 A:Molecule type: DNA  
 A:Residues: 1-144 <RES>  
 A:Cross-references: GB:M13207; NID:g181147; PIDN:AAA98768.1; PID:g181148  
 R:Cantrill, M.A.; Anderson, D.; Cerretti, D.P.; Price, V.; McKereghan, K.; Tushinski, R.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 6250-6254, 1985  
 A>Title: Cloning, sequence, and expression of a human granulocyte/macrophage colony-stim  
 A:Reference number: A25169; MUID:85298329; PMID:3898082  
 A:Accession: A25169  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <CAN>  
 A:Cross-references: GB:M1734; NID:g181149; PIDN:AAA52122.1; PID:g181150  
 R:Lee, F.; Yokota, T.; Otsuka, T.; Gemmell, L.; Larson, N.; Luh, J.; Aral, K.; Rennick,  
 Proc. Natl. Acad. Sci. U.S.A. 82, 4360-4364, 1985  
 A>Title: Isolation of cDNA for a human granulocyte-macrophage colony-stimulating factor  
 A:Reference number: A01853; MUID:85242684; PMID:3925454  
 A:Accession: A01853  
 A:Molecule type: mRNA  
 A:Residues: 1-144 <RES>  
 A:Cross-references: GB:M11220; NID:g183363; PIDN:AAA52578.1; PID:g183364  
 R:Hong, G.G.; Witek, J.S.; Temple, P.A.; Wilkens, K.M.; Leary, A.C.; Luxenberg, D.P.; Jo  
 Clark, S.C.  
 Science 248, 810-815, 1985  
 A>Title: Human GM-CSF: molecular cloning of the complementary DNA and purification of th  
 A:Reference number: A44175; MUID:85218749; PMID:3923663  
 A:Accession: A44175  
 A:Molecule type: mRNA  
 A:Residues: 1-116, 'T', 118-144 <MON>  
 A:Cross-references: GB:M10663; NID:g181145; PIDN:AAA52121.1; PID:g181146  
 A>Note: parts of this sequence, including the amino end of the mature protein, were conf  
 R:Men, D.Y.; Huang, B.R.; Cai, L.W.; Si, J.Y.  
 Acta Biochim. Biophys. Sin. 25, 651-655, 1993  
 A>Title: Amplification of human granulocyte-macrophage colony-stimulating factor cDNA wi  
 A:Accession: JCI090  
 A:Molecule type: protein  
 A:Residues: 18-21, 'C', 23-96, 'L', 98-144 <MEN>  
 C:Genetics:

A:Gene: GDB:CSF2  
 A:Cross-references: GDB:119812; OMIM:138960  
 A:Map position: 5q23.2-5q31.1  
 A:Introns: 53/3; 67/3; 109/3  
 C:Function:  
 A:Description: stimulates the differentiation and proliferation of hematopoietic prog  
 C:Superfamily: granulocyte-macrophage colony-stimulating factor  
 C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-144/Product: granulocyte-macrophage colony-stimulating factor #status experiment  
 F:44,54/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 18.08; Score 674; DB 1; Length 144;  
 Best Local Similarity 99.28; Pred. No. 7, 5e-31;  
 Matches 127; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 548 AAPARSPSPQPEHNAIOEARRLNLSRDPAEKNETVEYISEMFDQEPICLOTRL 607  
 Db 17 SAPARSPSPQPEHNAIOEARRLNLSRDPAEKNETVEYISEMFDQEPICLOTRL 76  
 Qy 608 ELVKGRLGSLTKLKPGLTMASHYKQCPPTPETSCTQITFESEKKNLKDPLVLPF 667  
 Db 77 ELVKGRLGSLTKLKPGLTMASHYKQCPPTPETSCTQITFESEKKNLKDPLVLPF 136

Qy 668 DCWEPEVOE 675  
 Db 137 DCWEPEVOE 144

RESULT 12  
 JCI387  
 epidermal growth factor receptor homolog precursor - rat  
 N:Alternate names: ErbB3 protein; HER3 protein  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 17-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 13-Nov-1998  
 C:Accession: JCI387  
 R:Hellier, N.J.; Kim, H.H.; Greaves, C.H.; Sterke, S.L.; Koland, J.G.  
 Gene 165, 279-284, 1995  
 A>Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protei  
 A:Reference number: JCI387; MUID:96096535; PMID:8522190  
 A:Accession: JCI387  
 A:Molecule type: mRNA  
 A:Residues: 1-1339 <HEU>  
 A:Cross-references: GB:U29339; NID:g915389; PID:g915390  
 A:Experimental source: Liver  
 A>Note: The authors translated the codon AAC for residue 369 as Thr and GAT for resid  
 C:Comment: This protein is a functional heregulin receptor that transduces signals to  
 C:Genetics:  
 A:Gene: ERBB3  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
 C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-1339/Product: epidermal growth factor homolog #status predicted <MAT>  
 F:640-659/Domain: transmembrane #status predicted <TM>  
 F:705-970/Domain: protein kinase homology (KIN)  
 F:713-721/Region: protein kinase ATP-binding motif  
 F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr)  
 F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr)

Query Match 17.7%; Score 662; DB 2; Length 1339;  
 Best Local Similarity 43.48; Pred. No. 4, 2e-29;  
 Matches 135; Conservative 44; Mismatches 108; Indels 24; Gaps 7;

Qy 17 GFLFLFFMIDRSVLAKELARGAA---STQVCTGTDKRLRPAESPETHLDMLRLVQGGC 73  
 Db 5 GTGVLVLCFL-----SLARGSEMGSQAVCGTGLNGLSVTDGADNOGYTLKVEKCE 57  
 Qy 74 VGVGNLETLPLPNASISFLQDIOEVQGYLVLIHNOVKRQVPLQRLRVKGTQLPEDNYAL 133  
 Db 58 VVGNNEIYLTGNADLSFLQWIREYAVYLVAMNESVYPLNLRVRYKTOYVGDGFAI 117  
 Qy 134 AVLVDGDPUNNTTPVVGASPGRLRLQLRLTEILKGVLIQNNPOLCYQDTILMKDIFH 193  
 Db 118 FVM-----LNYNT-----NSSHALRLQLTQLTEILISGVIYEKNDLCHMDITDMRDIV 168



Oy 668 DCWEPOE 675  
| | | | |  
Db 137 DCWEPAOK 144

Search completed: April 28, 2003, 13:42:22  
Job time : 32.3846 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:08 ; Search time 55.3846 seconds  
(without alignments)  
1660.081 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739  
Sequence: 1 MRAAPLLRAASLSIGFLF.....EPVQEGAPPPRAHHHHHH 630

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3739	100.0	690	22	AAE13109	Human HER500-hGM-C
2	3473.5	92.9	697	22	AAE13111	Human HER500-hGM-C
3	2972	79.5	535	22	AAE13108	Human HER500-fusio
4	2957.5	79.1	564	22	AAE13110	Human HER500-fusio
5	2573.5	68.8	919	21	AAE13103	Human HER-2/neu fu
6	2573.5	68.8	919	21	AAE13148	Her-2/neu extracel
7	2410.5	64.5	1255	21	AAE13148	Human heregulin 2
8	2410.5	64.5	1255	22	AAE12130	Human tyrosine kin
9	2410.5	64.5	1255	22	AAE12130	HER2 transgene pla
10	2410.5	64.5	1255	23	AAU74545	Human HER2 (ErbB2)

11	2405.5	64.3	1255	17	AAW01111	HER-2/neu protein.
12	2405.5	64.3	1255	20	AAW92406	Human HER-2/neu on
13	2405.5	64.3	1255	21	AAW21198	Human HER-2/neu pr
14	2405.5	64.3	1255	21	AAW84780	Amino acid sequenc
15	2405.5	64.3	1255	22	AAW85458	Human HER-2/neu pr
16	2405.5	64.3	1255	22	AAW88267	HER2/neu amino aci
17	2405.5	64.3	1255	23	AAE20407	Human Her-2 protel
18	2405.5	64.3	1255	23	AAE20479	Human Her-2/neu pr
19	2405.5	64.3	1255	23	AAW51143	Human Her-2/neu on
20	2405.5	64.3	1255	23	AAU77114	Human Her-2/neu po
21	2400.5	64.2	1433	14	AAW39568	Sequence of c-erbB
22	2334.5	62.4	1223	13	AAW98923	Human breast cance
23	2224	59.5	782	18	AAW19764	Her-2-GM-CSF immuno
24	2188.5	58.5	479	22	AAE13112	Human HER300-hGM-C
25	2130	57.0	926	23	AAW51153	Mouse Her-2/neu ex
26	2129	56.9	920	23	AAW51152	Mouse Her-2/neu ex
27	2092.5	56.0	1200	21	AAW21208	Human HER-2/neu pr
28	1961	52.4	1256	21	AAW21206	Mouse Her-2/neu pr
29	1961	52.4	1256	22	AAW21206	Amino acid sequenc
30	1961	52.4	1256	23	AAW51151	Mouse Her-2/neu on
31	1948	52.1	1256	21	AAW21199	Rat Her-2/neu prot
32	1948	52.1	1256	23	AAW51144	Rat Her-2/neu onco
33	1628.5	43.6	931	21	AAW44993	DC8cFv-erbB2EC fu
34	1612	43.1	645	22	AAW60408	Human ErbB2 oncopr
35	1612	43.1	645	22	AAW61593	Human ErbB2 extrac
36	1612	43.1	653	21	AAW21200	Extracellular HER-
37	1612	43.1	653	23	AAW51145	Human Her-2/neu on
38	1612	43.1	712	21	AAW21204	Human HER-2/neu fu
39	1612	43.1	712	23	AAW51149	Her-2/neu extracel
40	1587	42.4	289	22	AAE13120	Matute human HER-2
41	1562	41.8	624	21	AAW08222	Extracellular port
42	1553	41.5	419	23	AAE20348	Human truncated HE
43	1551	41.5	419	22	AAE09181	Human p68HHR-2 gen
44	1551	41.5	419	22	AAE09183	Human p68HHR-2 gen
45	1551	41.5	419	22	AAE09200	Human p68HHR-2 gen

#### ALIGNMENTS

RESULT 1	
AAE13109	AAE13109 standard; Protein: 690 AA.
ID	AAE13109
AC	AAE13109;
XX	
XX	28-JAN-2002 (first entry)
DE	
XX	Human HER500-hGM-CSF fusion protein construct.
XX	
KW	Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW	Immunostimulatory component; T-cell mediated immune response; DC;
KW	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW	PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW	membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW	HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW	HER500-hGM-CSF fusion protein.
XX	
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic.
XX	
PN	WO200174855-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US10515.
XX	
PR	30-MAR-2000; 2000US-193504P.
XX	
PA	(DEND-) DENDREON CORP.
XX	
PI	Laus R, Vidovic D, Graddis T;
XX	

DR	WPI: 2001-662965/76.
DR	N-PSDB; AAD21565.
PT	An Immunostimulatory fusion protein comprising the intracellular domain
PT	of HER-2 and an antigen elicits an immune response to the antigen and
PT	is useful for the treatment of associated cancer associated -
xx	
xx	
ps	Claim 7; Page 26; 59pp; English.
cc	
cc	The invention relates to immunostimulatory fusion proteins (IFP) and
cc	nucleic acid molecules encoding such proteins. The IFPs comprise a
cc	polypeptide antigen component and an immunostimulatory component derived
cc	from the intracellular domain of HER-2 protein which is effective to
cc	elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
cc	immune response to the antigen. IFP or superactivated dendritic cells
cc	are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
cc	associated with a particularly antigen. The present sequence is HER500
cc	hGM-CSF fusion protein construct which comprises human PAP
cc	signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
cc	signal sequence, mature HER-2 membrane distal extracellular and
cc	intracellular domains, an Ala Ala linker, a mature human granulocyte-
cc	macrophage colony stimulating factor (GM-CSF) sequence and a
cc	C-terminal tag.
xx	
xx	
SQ	Sequence 690 AA;
Query Match	100.0%; Score 3739; DB 22; Length 690;
Best Local Similarity	100.0%; Pred. No. 4,8e-226;
Matches 690; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 MRAAPLLAARASLSIGFLLEPFMLDRSVYLKELARGAASPOVCTGTMKRLRPAASPET 60
Db	1 MAAPLLAARAASLSIGFLLEPFMLDRSVYLKELARGAASPOVCTGTMKRLRPAASPET 60
OY	61 HLDMLRHLYQGCGVVGNLSELTLPRTNASLSFLDIOEYGVYLIANNOVROYPLQRRI 120
Db	61 HLDMLRHLYQGCGVVGNLSELTLPRTNASLSFLDIOEYGVYLIANNOVROYPLQRRI 120
OY	121 VAGTQLFEDNVNVALVLNDNDPLNTTPVTGASPGGIREQLDSLTETILGGVLIQNRPOL 180
Db	121 VAGTQLFEDNVNVALVLNDNDPLNTTPVTGASPGGIREQLDSLTETILGGVLIQNRPOL 180
OY	181 CYODITLMKDIDHNKNQALTLIDINRSACHPSCPMCKSGMOWSSSDCOSLTPTVCA 240
Db	181 CYODITLMKDIDHNKNQALTLIDINRSACHPSCPMCKSGMOWSSSDCOSLTPTVCA 240
OY	241 GGCAKCGKLPTDCCHEQCAAGCTGPKHSDDCLALFNHNSGICELCPALVYTNTDFES 300
Db	241 GGCAKCGKLPTDCCHEQCAAGCTGPKHSDDCLALFNHNSGICELCPALVYTNTDFES 300
OY	301 MPNPGGRATFGASCVTACPYNLTSTDVGSAGAMVHHRRSSSTRSGGDLTLGLEPSEE 360
Db	301 MPNPGGRATFGASCVTACPYNLTSTDVGSAGAMVHHRRSSSTRSGGDLTLGLEPSEE 360
OY	361 EAPRPRLAPSEAGSDVPDGDLGMAAKGLSLPHNDEPRDQRYSDDPVPLPSETDIGIV 420
Db	361 EAPRPRLAPSEAGSDVPDGDLGMAAKGLSLPHNDEPRDQRYSDDPVPLPSETDIGIV 420
OY	421 APVTSPPEYVNOPDVVRPPSPREGPLPARARPAGATLERKTISPNGNGVKDVFAG 480
Db	421 APVTSPPEYVNOPDVVRPPSPREGPLPARARPAGATLERKTISPNGNGVKDVFAG 480
OY	481 GAVENPEYLTPOGGAAPORHPRAPSPAFDNLYUMDOPREKGARSTFKGTPTANPEY 540
Db	481 GAVENPEYLTPOGGAAPORHPRAPSPAFDNLYUMDOPREKGARSTFKGTPTANPEY 540
OY	541 LGLDVPAAPAARSPTSPTDMEHVNAIOEARLNLISRDTAEMNETVEYISEMEDLOEP 600
Db	541 LGLDVPAAPAARSPTSPTDMEHVNAIOEARLNLISRDTAEMNETVEYISEMEDLOEP 600
OY	601 TGLDRLLEYLKYGLKSLTKLGRLTYMASSHKKOCPRPPESSCATQTTFESFEKNLMD 660
Db	601 TGLDRLLEYLKYGLKSLTKLGRLTYMASSHKKOCPRPPESSCATQTTFESFEKNLMD 660

QY	661	FLTVLPEDCMEPVQEGAPPPAAHHNNH	690
DB	661	FLTVLPEDCMEPVQEGAPPPAAHHNNH	690
		RESULT 2	
		AAE13111	
ID	AAE13111	standard; Protein; 697 AA.	
AC	AAE13111;		
XX			
DT	28-JAN-2002	(first entry)	
DE			
XX			
KM	Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide:		
KM	Immunostimulatory fusion protein; IFP; antigen component; therapy;		
KM	immunostimulatory component; T-cell mediated immune response; DC;		
KM	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;		
KM	PAP protein; Ala Arg linker; membrane distal extracellular domain;		
KM	membrane distal intracellular domain; C-terminal tag; human; GM-CSF;		
KM	HER-2 protein; granulocyte-macrophage colony stimulating factor;		
KM	ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.		
XX			
OS	Chimeric - Homo sapiens.		
OS	Chimeric - Rattus norvegicus.		
XX	Chimeric - Unidentified.		
PN	WO200174855-A2.		
PD	11-OCT-2001.		
PF	30-MAR-2001; 2001WO-US10515.		
PR	30-MAR-2000; 2000US-193504P.		
PA	(DEND-) DENDREON CORP.		
PI	Laus R, Vladovic D, Graddis T;		
DR	WPI: 2001-662965/76.		
PS	N-PSDB; MAD21567.		
CC	An immunostimulatory fusion protein comprising the intracellular domain		
CC	of HER-2 and an antigen elicits an immune response to the antigen and		
CC	is useful for the treatment of associated cancer associated -		
CC	Claim 7; Page 27; 59pp; English.		
CC	The invention relates to immunostimulatory fusion proteins (IFP) and		
CC	nucleic acid molecules encoding such proteins. The IFPs comprise a		
CC	polypeptide antigen component and an immunostimulatory component derived		
CC	from the intracellular domain of HER-2 protein which is effective to		
CC	elicit a protective dendritic cell (DC)-induced T-cell mediated cellular		
CC	immune response to the antigen. IFP or superactivated dendritic cells		
CC	are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer		
CC	associated with a particularly antigen. The present sequence is HER500		
CC	rGM-CSF fusion protein construct which comprises human PAP		
CC	signal sequence, mature PAP protein, an Ala Arg linker, human HBR-2		
CC	signal sequence, mature HER-2 membrane distal extracellular domain,		
CC	an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,		
CC	HER-2 membrane distal intracellular domain, an Ala Ala linker, a mature		
CC	rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence		
CC	and a C-terminal tag.		
SO	Sequence 697 AA:		
Query Match	92.9%; Score 3473.5; DB 22; Length 697;		
Best Local Similarity	91.7%; Pred. NO. 2.le-209;		
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2			
1	MRAPELLARAASLSIGLFFLLFWLNDSVLAKELRGAASIVGCTGTDMKRLRPASPT 60		



XX	AAEL3110	standard: Protein; 564 AA.	
XX	AAEL3110;		
XX	28-JAN-2002	(first entry)	
XX	Human HER500 fusion protein construct comprising OVA-derived octapeptide.		
XX	Immunostimulatory fusion protein; IFP; antigen component; therapy;		
XX	Immunostimulatory component; T-cell mediated immune response; DC;		
XX	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;		
XX	PAP protein; Ala Arg linker; membrane distal extracellular domain;		
XX	membrane distal intracellular domain; C-terminal tag; human; OVA;		
XX	HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.		
XX	Chimeric - Homo sapiens.		
XX	Chimeric - Unidentified.		
XX	MO200174855-A2..		
XX	11-OCT-2001.		
XX	30-MAR-2001; 2001WO-US10515.		
XX	30-MAR-2000; 2000US-193504P.		
XX	(DEND-) DENDREON CORP.		
XX	Laus R, Vidovic D, Graddis T;		
XX	WPI; 2001-662965/76.		
XX	N-PSDB; AAD21566.		
XX	An immunostimulatory fusion protein comprising the intracellular domain		
XX	of HER-2 and an antigen elicits an immune response to the antigen and		
XX	is useful for the treatment of associated cancer associated -		
XX	Claim 7; Page 26; 59p; English.		
XX	The invention relates to immunostimulatory fusion proteins (IFP) and		
XX	nucleic acid molecules encoding such proteins. The IFPs comprise a		
XX	polypeptide antigen component and an immunostimulatory component derived		
XX	from the intracellular domain of HER-2 protein which is effective to		
XX	elicit a protective dendritic cell (DC)-induced T-cell mediated cellular		
XX	immune response to the antigen. IFP or superactivated dendritic cells		
XX	are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer		
XX	associated with a particularly antigen. The present sequence is HER500		
XX	fusion protein construct which comprises human PAP signal		
XX	sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal		
XX	sequence, mature HER-2 membrane distal extracellular domain, an		
XX	Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,		
XX	HER-2 membrane distal intracellular domain and a C-terminal tag.		
XX	Sequence 564 AA:		
XX	Query Match 79.1%; Score 2957.5; DB 22; Length 564;		
XX	Best Local Similarity 98.4%; Pred. No. 3.4e-177;		
XX	Matches 549; Conservative 0; Mismatches 0; Indels 9; Gaps 1		
XX	1 MKAAPLLARAASTLSIGFLFLFPMWDRSVLAKKELARGAASOVYCGTDMKRLPASPET 60		
XX	1 MRAAPLLARAASTLSIGFLFLFPMWDRSVLAKKELARGAASOVYCGTDMKRLPASPET 60		
XX	61 HLDMLRHLYOGQGVVQGNLELTYPFTNASTLSFLQDIOEVQGVLLAHNOVRQPLDRLRI 120		
XX	61 HLDMLRHLYOGQGVVQGNLELTYPFTNASTLSFLQDIOEVQGVLLAHNOVRQPLDRLRI 120		
XX	121 VAGTDLFEDNVALAVLNDGDP LANTPTVTGASPGGLREQLASLFTLKGCVLIQRNPL 180		

Db	121	VRGQLEFEDNYALAVLNDGDRPLNNTTPEVTGASPGGLRELDLRSLTEILKGVLQIORNPOL	180
Qy	181	CYODTILMKDIDFHKNNOALTLTDITNRSRACHPCSPMKSGSRKGESEEDCOSTIRVCA	240
Db	181	CYODTILMKDIDFHKNNOALTLTDITNRSRACHPCSPMKSGSRKGESEEDCOSTIRVCA	240
Qy	241	GGCARCKRPLTDDCHDQCAAGCTGPRKNSDCLAHNHSIGICLHCPALVYNTOTFEES	3000
Db	241	GGCARCKRPLTDDCHDQCAAGCTGPRKNSDCLAHNHSIGICLHCPALVYNTOTFEES	3000
Qy	301	MPNPEGRTFCAVCATCAQPVYVLTSDVGS-----GAGGWNHHNRSSSTRSGGDL	351
Db	301	MPNPEGRTFCAVCATCAQPVYVLTSDVGSASIIINERKLGAAGWNHNRSSSTRSGGDL	360
Qy	352	TLGLPSESEEARSPRLAPSEGAGSDVFDGDLGMAAGLQSLPHNDSPLQRTSEDTVP	411
Db	361	TLGLPSESEEARSPRLAPSEGAGSDVFDGDLGMAAGLQSLPHNDSPLQRTSEDTVP	420
Qy	412	LPSETDGVAAVLTCSPQPEYVNOQDVVRPPSPREGRPLAARAPAGATLEBAKTLSPKNG	471
Db	421	LPSETDGVAAVLTCSPQPEYVNOQDVVRPPSPREGRPLAARAPAGATLEBAKTLSPKNG	480
Qy	472	VVKDVFARFAGAVENPEYITPQGGAAAPRPHPPAPSPAFDNLVYWDQDPPERGAAPSTFKG	531
Db	481	VVKDVFARFAGAVENPEYITPQGGAAAPRPHPPAPSPAFDNLVYWDQDPPERGAAPSTFKG	540
Qy	532	TPTAENPEYIGLDVPAAA 549	
Db	541	TPTAENPEYIGLDVPAAA 558	

RESULT 5	
AAB21203	
ID	AAB21203 standard; protein; 919 AA.
XX	
AC	AAB21203;
XX	
DT	12-JAN-2001 (first entry)
XX	
DE	Human HER-2/neu fusion protein.
XX	
KW	Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW	breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW	colon cancer; fusion protein.
XX	
OS	Homo sapiens.
OS	Synthetic.
PN	WO20004489-A1.
XX	
PD	03-AUG-2000.
XX	
FP	28-JAN-2000; 2000MO-US02164.
XX	
PR	29-JAN-1999; 99US-0117976.
XX	
PA	(CORI-) CORIYA CORP.
PA	(SMIK ) SMITHKLINE BEECHAM.
XX	
PI	Cheever MA, Gheysen D;
XX	
DR	WPI: 2000-505976/45.
DR	
PT	HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT	useful for vaccinating against breast, ovarian, colon, lung and
PT	prostate cancers -
XX	
PS	Claim 2; Fig 12; 128pp; English.
XX	
CC	The present sequence is a fusion protein comprising the extracellular
CC	domain and the phosphorylation domain of the human HER-2/neu protein.
CC	HER-2/neu is a member of the tyrosine kinase family of receptor-like
CC	glycoproteins and shows homology to the epidermal growth factor receptor



CC (EGFR). It probably plays a part in cell growth and/or differentiation.  
 CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used  
 CC to treat or prevent cancer by eliciting or enhancing an immune response  
 CC to the HER-2/neu protein. They may be used to treat malignancies such as  
 CC breast, ovarian, colon, lung and prostate cancers, and may be used as an  
 CC antigen to vaccinate against these neoplasias.

XX Sequence 919 AA;

Query Match 68.8%; Score 2573.5; DB 21; Length 919;

Best Local Similarity 56.4%; Pred. No. 7, Le-153;

Matches 509; Conservative 0; Mismatches 3; Indels 391; Gaps 1;

```

QY 35 LARGAASVTCGTDMKRLPASPETHLMDLRHLYGCGVQVQGNLELYLPTNASTLSLQ 94
D 16 LPPGASSTQVTCGTDMKRLPASPETHLMDLRHLYGCGVQVQGNLELYLPTNASTLSLQ 75
QY 95 DIOEVQGYVLIANQVQVPLQRLIRIVRGTLFEDNYVALVNDGDPINNTPTVYGASPG 154
D 76 DIOEVQGYVLIANQVQVPLQRLIRIVRGTLFEDNYVALVNDGDPINNTPTVYGASPG 135
QY 155 GLRELQLSLLEIKGGLVLIQNRNQLCYQDTILMKDIFHKNNQALATLIDNRSRACHPC 214
D 136 GLRELQLSLLEIKGGLVLIQNRNQLCYQDTILMKDIFHKNNQALATLIDNRSRACHPC 195
QY 215 SPMCKSGKMGESSEDCOSLTRVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLAC 274
D 196 SPMCKSGKMGESSEDCOSLTRVCAGGACARCKGPLPTDCCHCCAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVYNTDTEFSMPNPGRYTFGASCVTACPYNTLSTDVGS---- 329
D 256 LHFHNSGICELHCPALVYNTDTEFSMPNPGRYTFGASCVTACPYNTLSTDVGSCTLVC 315
QY 330 ----- 329
D 316 PLHNOVTAEADGTORCEKSPCARVCYGLMEHLREYRAVTSANIQEACCKIIFGSLA 375
QY 330 ----- 329
D 376 FLPSFDGDPASNTAPLQPELOVFELEETGYIISAMPDLSLPQNLQYIRRI 435
QY 330 ----- 329
D 436 LHNAGASLTLLGLISWLGSLRLSRLSGSLALIHNNTHLCEVHVPMDQLFRNPHQALLH 495
QY 330 ----- 329
D 496 TANPDEBCVEGLACHQDLARHGCMGPGPTQCVCNCSQFLRGQCEVCRLVGLPREYV 555
QY 330 ----- 329
D 556 NARHCLPCHPECPQNSVTCFGEADQCVACANHKDPFCVARGSPGVKDLISWYIMK 615
QY 330 ----- 329
D 616 FPDEGACQPCPINCTHSCVLDLDDKGCFAEORASPLTSQNEDLGASPLDSTFYRSLLED 675
QY 330 -----GAGGVVHHRHSSRSRSGGDLTILEPSEEBAP 735
D 676 DDMKDLDAEYLVLPQGFECPRDPAFGAGVHHRHSSRSRSGGDLTILEPSEEBAP 735
QY 364 RSLPLASGAGSDVFDGDLGWAAGKGLQSLPTHDPSPLOQRYSEDPVLPSETDGYVAPL 423
D 736 RSLPLASGAGSDVFDGDLGWAAGKGLQSLPTHDPSPLOQRYSEDPVLPSETDGYVAPL 795
QY 424 TCSQPEYVNPQVPRPSPREGRPLPAARAGATLERAKTLSPGKNGVAVDFAFGAV 483
D 796 TCSQPEYVNPQVPRPSPREGRPLPAARAGATLERAKTLSPGKNGVAVDFAFGAV 855
QY 484 ENPEYTLPPGGGAAPQHPHPPAFSPAFDNLYWDDPPERGAAPSTFKGTPTAENPEYIGL 543
D 856 ENPEYTLPPGGGAAPQHPHPPAFSPAFDNLYWDDPPERGAAPSTFKGTPTAENPEYIGL 915

```

QY 544 DVP 546  
 D 916 DVP 918

RESULT 6

ID AAM51148

AA51148 standard; Protein: 919 AA.

AC AAM51148;

DT 17-JUN-2002 (first entry)

DE Her-2/neu extracellular domain-phosphorylation domain fusion.

KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;

KW Tyrosine kinase; receptor; c-erbB2; gene therapy.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Domain 1..653

FT Domain /note="extracellular domain"

FT Domain 654..919

FT Domain /note="phosphorylation domain"

PD WO200212341-A2.

PD 14-FEB-2002.

PF 03-AUG-2001; 2001WO-US24283.

PF 03-AUG-2000; 2000US-0632507.

PA (CORI-) CORIYA CORP.

PA (SMIR ) SMITHKLINE BEECHAM BIOLOGICALS.

PI Cheever MA, Gheysen D;

PI WPI: 2002-241743/29.

PS Claim 2; Fig 12; 141pp; English.

XX The present sequence is that of a fusion protein between the  
 XX extracellular domain and phosphorylation domain of human Her-2/neu  
 XX (see AAM51143), an oncogenic self-protein and target for anti-cancer  
 XX vaccines. The fusion protein can be obtained by recombinant DNA  
 XX methods. Her-2/neu overexpression correlates with a poor prognosis  
 XX in breast and ovarian cancers. The invention provides Her-2/neu  
 XX vaccines comprising the fusion proteins or nucleic acid molecules.  
 XX In preferred fusion proteins, the extracellular domain of a  
 XX Her-2/neu protein is fused to a Her-2/neu intracellular domain or  
 XX phosphorylation domain (or its deltaD fragment). An immune  
 XX response to Her-2/neu protein is elicited or enhanced by  
 XX administering the fusion protein in the form of a vaccine, or by  
 XX transfecting cells of an animal ex vivo with a nucleic acid  
 XX encoding the fusion protein, and delivering the transfected cells  
 XX to the animal. The fusion proteins, nucleic acids, and isolated  
 XX specific T-cells are useful for inhibiting the development of a  
 XX cancer, especially breast, ovarian, colon, lung or prostate cancer  
 XX in a patient. T cells that specifically react with a Her-2/neu  
 XX fusion protein can be used to remove tumour cells from a sample in  
 XX order to inhibit the development of cancer in a patient.

Query Match 68.8%; Score 2573.5; DB 21; Length 919;

Best Local Similarity 56.4%; Pred. No. 7.1e-153;  
Matches 509; Conservative 0; Mismatches 3; Indels 391; Gaps 1;

```
OY 35 LARGASTOYCTGDMKLRIPASPEHLLDMLRLHYOCQOVVGNLTETLYPTNASTSFLQ 94
DB 16 LPPGASTOYCTGDMKLRIPASPEHLLDMLRLHYOCQOVVGNLTETLYPTNASTSFLQ 75
OY 95 DIEVOGVYLIANNOVROVPLQRLRYRGTQLEEDNVALVLDNGDPLNNTPTVTGASPG 154
DB 76 DIEVOGVYLIANNOVROVPLQRLRYRGTQLEEDNVALVLDNGDPLNNTPTVTGASPG 135
OY 155 GLAEQLRLSTLEIKGCVLIQIRNPOLCYODTIIMKDIFFHKNNOIALTLIDTNRSRACHPC 214
DB 136 GLAEQLRLSTLEIKGCVLIQIRNPOLCYODTIIMKDIFFHKNNOIALTLIDTNRSRACHPC 195
OY 215 SPACKSGRCGSESEDCOSLTRTYCAGGACRCRGPPLPTDCHECCAGCTGPKHSDDLAC 274
DB 196 SPACKSGRCGSESEDCOSLTRTYCAGGACRCRGPPLPTDCHECCAGCTGPKHSDDLAC 255
OY 275 LRFNHSICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPNYULSTDVGS---- 329
DB 256 LRFNHSICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPNYULSTDVGSCTLYC 315
OY 330 ----- 329
DB 316 PLANOEVTAEDGTORCEKSKPCARVCYGLGMEHLREAVTSAPIOEFAGCKKIFGSLA 375
OY 330 ----- 329
DB 376 FLRFSPGDPAASMTAPLQRPQLOVFETLEITGVLVISAMPDLSLPDLSVONLQVIRGRI 455
OY 330 ----- 329
DB 436 LHNAGVSLTQIGISWLGSLRLSELGSLALIHNTHTLCFVHTVPWDQLFRNHQALLH 495
OY 330 ----- 329
DB 496 TANRPEDECVGEGILACHQLCARGHCWGPPTQCVNCSQPLRGQECVECHVLOGLPREV 555
OY 330 ----- 329
DB 556 NARHCLPCHPEQOPONGSVTCFPEADQCVACAHYKDPFCVARCPSGVKPRDLSYMPIMK 615
OY 330 ----- 329
DB 616 FPDEGACOPPCINCHSCVLDKGCPRABORASPLTSQNEIDLGPASPLDSTFYRSLLD 675
OY 330 ----- 329
DB 676 DMGDLVDAEEYLVPOOGFCPPAPACAGMHHRRHSSSTRSGGDLTLGLEPSEEEAP 735
OY 364 RSLPLASEGASDVFDGDLGMAAKGLSLPTHDPSPLQRYSDPVPPLPSETDGYVAPL 423
DB 736 RSLPLASEGASDVFDGDLGMAAKGLSLPTHDPSPLQRYSDPVPPLPSETDGYVAPL 795
OY 424 TCSPOREYVQPPVQPPSPRREGPLPAARPAATLELRATLSLSPGKGVYKDVFAFGAV 483
DB 796 TCSPOREYVQPPVQPPSPRREGPLPAARPAATLELRATLSLSPGKGVYKDVFAFGAV 855
OY 484 ENBEYLTPOGGAAPQHPAPAFSPAFDNLTYMDDPPERAGAPSTFGTPTAENPEYGL 543
DB 856 ENBEYLTPOGGAAPQHPAPAFSPAFDNLTYMDDPPERAGAPSTFGTPTAENPEYGL 915
OY 544 DVP 546
DB 916 DVP 918
```

```
XX 10-AUG-2000 (first entry)
DT Human heregulin 2 (Her2).
XX
DE Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
KW self-protein; cancer; breast cancer; prostate cancer;
KW cell-associated peptide antigen; foreign epitope.
XX
OS Homo sapiens.
XX
FH Location/Qualifiers
FT 1..173
FT /label= N-terminal
FT /note= "mature polypeptide"
FT 5..25
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 59..73
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 103..117
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 149..163
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 174..323
FT /label= Cysteine_rich_domain
FT 210..224
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 250..264
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 324..483
FT /label= ligand_binding_domain
FT 325..339
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 369..383
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 465..479
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 484..623
FT /label= Cysteine_rich_domain
FT 579..593
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 624..654
FT /label= Transmembrane_domain
FT 632..652
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 653..667
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 655..1010
FT /label= Tyrosine_kinase_domain
FT 661..675
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 695..709
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 710..730
FT /label= insertion_region
FT /note= "suitable for foreign epitope insertion"
FT 1011..1235
FT /label= C-terminal_domain
XX
```





DE HER2 transgene plasmid construct encoded protein.  
 XX  
 KW Human: HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;  
 KW antibody.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN W0200100244-A2.  
 XX  
 PD 04-JAN-2001.  
 XX  
 PF 23-JUN-2000; 2000MO-US17229.  
 XX  
 PR 25-JUN-1999; 99US-0141316.  
 PR 16-MAR-2000; 2000US-0189844.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Erickson S, Schwall R;  
 XX  
 DR WPI: 2001-061962/07.  
 DR N-PSDB: AAF24297.  
 XX  
 XX Treating tumors, particularly breast cancers, which overexpress an ErbB  
 PT receptor and does not respond to an anti-ErbB antibody, comprises  
 PT conjugating the antibody to a maytansinoid -  
 XX  
 PS Example 3; Fig 4; 92pp: English.  
 XX  
 CC The present invention provides a method of treating cancer by  
 CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In  
 CC particular, the antibody is directed against ErbB2 (also known as HER2  
 CC and p185neu). The method is particularly useful in the treatment of  
 CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,  
 CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.  
 CC  
 XX  
 SO Sequence 1255 AA:  
 Query Match 64.5%; Score 2410.5; DB 22; Length 1255;  
 Best Local Similarity 41.2%; Pred. No. 1,7e-142;  
 Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

Db 436 LHNCAVSLTLOGIGISWLGSLRLSRLSGLALIHNNTHLCEVHTVPMWDLFRNPHOALLH 495  
 Qy 330 ----- 329  
 Db 496 TANRDECEVGEGLACHOLCARGHCWGPPTQVCNCSQFLRGQCEVCRLQGLPREV 555  
 Qy 330 ----- 329  
 Db 556 NARHCLPCHPECOQONGSVTCGPEADQVACAHYKDPFCVARGCPGKPDLSYMPIWK 615  
 Qy 330 ----- 329  
 Db 616 FPDEGACQPCPINCTHSCVDLDKCPAEQBRASPLTSTVSAVGLLVVLGVVFGILI 675  
 Qy 330 ----- 329  
 Db 676 KRROQKIRKYMRLLOETELVEPLTPSGAMPNQAOMRLKETELKVKVIGSGAFGTVY 735  
 Qy 330 ----- 329  
 Db 736 KGIWIPDGENKIPVAIVLRENTSPRANKETLDEAVYMAVGSPYVSRLLGICLTSTVO 795  
 Qy 330 ----- 329  
 Db 796 LVTQLMPYGLLDHVRENRRGLSODLLNMCQIAKGSYLEDVRLVHRDLAARNVLVS 855  
 Qy 330 ----- 329  
 Db 856 PNHVKITDFGLARLDIDETEHADGKRVPIKMALLESILRRRTHQSDVMSYGVTWEL 915  
 Qy 330 ----- 329  
 Db 916 MTFGAKPTDGIPIAREIPDLLEKGERLPQPPICITDVIYIMVKCMIDSECRPRELYSE 975  
 Qy 330 ----- 329  
 Db 976 FSRMARDPQRFVYIGNEDLPASPLDSTFYRSILEDMDGDLVDAEYLYPQGFECDDP 1035  
 Qy 330 ----- 329  
 Db 1036 APGAGMWHHRSSSTSGGDDLTGLSPSEEARPSPLAPSEAGSDVDEGDMKMAA 1095  
 Qy 388 KGLGSLPTNHDSPLORYEEDPTVPLPSETDGVAVARLTSQPEVYNQDPVRPQSPREG 447  
 Db 1096 KGLGSLPTNHDSPLORYEEDPTVPLPSETDGVAVARLTSQPEVYNQDPVRPQSPREG 1155  
 Qy 448 PLPAARPAATLERAKTSLSPGKNGVYKDVAFAGAVENPEYLTPOGGAAPORHPAPESP 507  
 Db 1156 PLPAARPAATLERAKTSLSPGKNGVYKDVAFAGAVENPEYLTPOGGAAPORHPAPESP 1215  
 Qy 508 AFDNLVYWDODPPEKAPSPFTFKGPTAENPEYLGLDVP 546  
 Db 1216 AFDNLVYWDODPPEKAPSPFTFKGPTAENPEYLGLDVP 1254  
 RESULT 10  
 AAU74545  
 ID AAU74545 standard; Protein: 1255 AA.  
 XX  
 AC AAU74545;  
 XX  
 DT 23-APR-2002 (first entry)  
 XX  
 DE Human HER2 (ErbB2) polypeptide.  
 KW  
 KW Human: HER2; ErbB; epidermal growth factor receptor; receptor;  
 KW anti-ErbB antibody-maytansinoid conjugate; cancer; tumour; breast; ovary;  
 KW stomach; endometrium; salivary gland; lung; kidney; colon; colorectum;  
 KW thyroid; pancreas; prostate; bladder; ErbB2; neuronal disorder;  
 KW glial disorder; astrocytal disorder; hypothalamic disorder;  
 KW glandular disorder; macropinagal disorder; epithelial disorder;  
 KW stromal disorder; blastocoeleic disorder; inflammatory disorder;  
 KW angiogenic disorder; immunological disorder;

XX Homo sapiens.  
OS  
XX US2002001587-A1.  
PN  
XX  
PD 03-JAN-2002.  
XX  
PF 16-MAR-2001; 2001US-0811123.  
XX  
PR 16-MAR-2000; 2000US-189844P.  
PR 05-OCT-2000; 2000US-238327P.  
XX  
PA (ERIC/) ERICKSON S.  
PA (SCHM/) SCHMALL R.  
PA (SLIM/) SLIMKOWSKI M.  
XX  
PI Erickson S, Schwall R, Slilkowski M;  
XX  
DR WPI: 2002-163686/21.  
DR N-PSDB: ABR14058.  
XX  
XX  
PT Treating tumour characterised by overexpression of epidermal growth  
PT factor receptor, ErbB or cancer in mammal, comprises administering  
PT anti-ErbB antibody-maytansinoid conjugate to the mammal  
XX  
PS Example 3; Fig 7; 93pp; English.  
XX  
XX The invention relates to treating a tumour in a mammal, where the tumour  
CC is characterised by the overexpression of an epidermal growth factor  
CC receptor (ErbB) and does not respond or responds poorly, to treatment  
CC with an anti-ErbB antibody, comprising administering to the mammal an  
CC anti-ErbB antibody-maytansinoid conjugate. The method is useful for  
CC treating cancer or tumours of the breast, ovary, stomach, endometrium,  
CC salivary gland, lung, kidney, colon, colorectum, thyroid, pancreas,  
CC prostate and bladder, preferably breast cancer. The breast cancer is a  
CC metastatic breast cancer or an aggressive form of metastatic breast  
CC cancer which overexpresses ErbB2. The method is also useful for treating  
CC neuronal, glial, astrocytal, hypochalamic, glandular, macrophagal,  
CC epithelial, stromal, blastocellic, inflammatory, angiolary, and  
CC immunological disorders. This sequence represents the human HER2 (ErbB2)  
CC polypeptide of the invention.  
XX  
SO Sequence 1255 AA:

Query Match 64.5%; Score 2410.5; DB 23; Length 1255;  
Best Local Similarity 41.2%; Pred. No. 1,7e-142;  
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LANGAASCTGCTGDMKRLRPASPEHLDMLRHLXGCGVVOGMLLETYLPNTASLSFLQ 94  
DB 16 LPPGAASCTGCTGDMKRLRPASPEHLDMLRHLXGCGVVOGMLLETYLPNTASLSFLQ 75  
QY 95 DIOEVGVYLIANOVQVFLQRLIRYGTQLFEDNVALAVLDNGDPLNNTTPVTGASPG 154  
DB 76 DIOEVGVYLIANOVQVFLQRLIRYGTQLFEDNVALAVLDNGDPLNNTTPVTGASPG 135  
QY 155 GLRELQRLSTLEIKGGLVLIQNRNQLCYQDTILMKDIFHKNNQDLATLIDTNSRACHPC 214  
DB 136 GLRELQRLSTLEIKGGLVLIQNRNQLCYQDTILMKDIFHKNNQDLATLIDTNSRACHPC 195  
QY 215 SPMCKSGRCGSESEDCOSILRTVYACAGGACRCAGPLPTDCCHQCAAGCGPKHSDCLAC 274  
DB 196 SPMCKSGRCGSESEDCOSILRTVYACAGGACRCAGPLPTDCCHQCAAGCGPKHSDCLAC 255  
QY 275 LHNHNSGICGLHCPALVTYNTDFESMPNPEGRTGASCVTACPYVYSTDGS----- 329  
DB 256 LHNHNSGICGLHCPALVTYNTDFESMPNPEGRTGASCVTACPYVYSTDGSCTLYC 315  
QY 330 ----- 329  
DB 316 PLHNOEYTAEDGTORCEKCKPCARVCYGLGMEHLREVRVTSANTIOEFAGCKKIEGSLA 375  
QY 330 ----- 329

DB 376 FLPEFDDGASNTAPLQPOLOVETLEITGYLIAMPDSLPLDSVFQNLQVINGRI 435  
QY 330 ----- 329  
DB 436 LHNCAVSLTLQGLGISWGLRLSRELQSGALALHNHNTLHCFVHTVPWDQLFRNPQALLH 495  
QY 330 ----- 329  
DB 496 TANRPEDECVGEGLAGHQLCARGHCMPGPTQVCNCSQFLRGQCEVECRVLOGLPREYV 555  
QY 330 ----- 329  
DB 556 NARHCLPHECPQPNQNSVTCFGEADQVACAHYKDPFPCVARGPSGVKPDLSYPMWK 615  
QY 330 ----- 329  
DB 616 FPDEGACQPPINCTHSCVDLDDKGPARGASPLTSYSAVVGILLVYLVGVPBILI 675  
QY 330 ----- 329  
DB 676 KRRQOKIRKYTMRLLOETELVEPLTPSGAMPNQAMRLKETELRKVKVSGAGFTVY 735  
QY 330 ----- 329  
DB 736 KGWIIPDENVKIPVAIKVLENTSPKANKEILDEAVVMAGVSPYVSRLLGICLTSTVQ 795  
QY 330 ----- 329  
DB 796 LVYQMLMPYGLLDHVRNRRGLSGQDLNMCQIAKMSVLEDRVLRHDLAARNVLYKS 855  
QY 330 ----- 329  
DB 856 PNHYKITDFGLARLDIDETEHADGKVPDKMALESILRRRPTHOSDVWYGVYVWEL 915  
QY 330 ----- 329  
DB 916 MTEGAKPYDGIPIAREIPDLLEKGERLPQPICTIDYIMVWKMDSECHPRELYSE 975  
QY 330 ----- 329  
DB 976 FSRMARDPQRFVYIQNEDLPASPLDSTFYRSLEDDOMGDLVDAEYLVPOQGFPCDP 1035  
QY 330 --GAGGVVHRRHSSSTRSGGDLTLCLEPSEEPAPSPAPSGACSDVYDGLGCAA 387  
DB 1036 APGAGGVVHRRHSSSTRSGGDLTLCLEPSEEPAPSPAPSGACSDVYDGLGCAA 1095  
QY 388 KGLQSLPTHPSPLORYSEDPVYPLPSETDGYVAPLCSPOPEYVNOPDVRRPOPSPREG 447  
DB 1096 KGLQSLPTHPSPLORYSEDPVYPLPSETDGYVAPLCSPOPEYVNOPDVRRPOPSPREG 1155  
QY 448 PLPARAPGATLERAKTILSPGKNGVVDVAFGAVENPEYLPQGGAPQHPPPAFSP 507  
DB 1156 PLPARAPGATLERAKTILSPGKNGVVDVAFGAVENPEYLPQGGAPQHPPPAFSP 1215  
QY 508 AFQNLVYWDQDPPEKAPSPSTFKGTPTAENPEYLGIDVP 546  
DB 1216 AFQNLVYWDQDPPEKAPSPSTFKGTPTAENPEYLGIDVP 1254

RESULT 11  
AAW01111  
ID AAW01111 standard; Protein; 1255 AA.  
XX AAW01111:  
AC  
XX  
DT 01-JAN-1997 (first entry)  
XX  
DE HER-2/neu protein.  
XX  
KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;  
KW breast cancer; ovary cancer; colon cancer; lung cancer;  
KW prostate cancer; immunisation; tumour; vaccine; vector.

XX OS Homo sapiens. Location/Qualifiers  
XX FH Key 676..1255  
FT Domain /label=intracellular\_domain  
FT /note="claimed domain, useful for immunisation"  
XX PN M09630514-A1.  
XX PD 03-OCT-1996.  
XX PE 28-MAR-1996; 96WO-US01689.  
XX PR 31-MAR-1995; 95US-0414417.  
XX PA (UNIW ) UNIV WASHINGTON.  
XX PI Cheever MA, Distis ML;  
XX DR WPI: 1996-455361/45.  
XX DR N-PSDB; AAT40739.  
XX PT DNA encoding HER-2/neu poly:peptide(s) - used for prevention or  
XX PT treatment of malignancies with which the HER-2/neu oncogene is  
XX PT associated  
XX PS Claim 2; Page 56-61; 71pp; English.  
XX CC Human HER-2/neu protein (AAM01111), also called p185 or c-erbB2, is  
XX CC the product of the HER-2/neu oncogene (see also AAT40739). The  
XX CC protein is over-expressed in various cancers, including breast,  
XX CC ovarian, colon, lung and prostate. The intracellular domain of the  
XX CC protein can be used to immunise an animal against a malignancy with  
XX CC which the oncogene is associated. The polypeptide can be produced  
XX CC in transformed host cells for use in immunisation. Alternatively,  
XX CC animal cells are transfected in vivo or ex vivo with a viral vector  
XX CC that directs expression of the polypeptide.  
XX SO Sequence 1255 AA;  
Query Match 64.3%; Score 2405.5; DB 17; Length 1255;  
Best Local Similarity 41.1%; Pred. No. 3.5e-142;  
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;  
QY 35 LARGAASOVCTGTCTMKTLRASPEETHDMLRHLYQCGQVVOGNLELYLPNASTLFLQ 94  
DB 16 LPPGAASOVCTGTCTMKTLRASPEETHDMLRHLYQCGQVVOGNLELYLPNASTLFLQ 75  
QY 95 DIOEVQGVLLAHNOVROVPLQRLRIVRGTOLEEDNYALAVLNDGDPPLNTPVTGASPG 154  
DB 76 DIOEVQGVLLAHNOVROVPLQRLRIVRGTOLEEDNYALAVLNDGDPPLNTPVTGASPG 135  
QY 155 GLRELQLRSLTEILKGVLIQNRNOLCYODTILMKDIFHKNNQALTLIDTNRSRACHPC 214  
DB 136 GLRELQLRSLTEILKGVLIQNRNOLCYODTILMKDIFHKNNQALTLIDTNRSRACHPC 195  
QY 215 SPMCKSGKMGWSSSEDCSLTFTVAGCARGKGPLPTDCCHEGCAAGCTGKRHSDCLAC 274  
DB 196 SPMCKSGKMGWSSSEDCSLTFTVAGCARGKGPLPTDCCHEGCAAGCTGKRHSDCLAC 255  
QY 275 LAFNHSIGICELCPALVYNTDTFESMPNPEGRTYFGASCYACVYNTLSTVGS----- 329  
DB 256 LAFNHSIGICELCPALVYNTDTFESMPNPEGRTYFGASCYACVYNTLSTVGSCTLVLC 315  
QY 330 ----- 329  
DB 316 PLHNOEVAEDGTQRCCKSPCARVYCYGLMEHLREAVATYSAIOEFAGCKITFGSLA 375  
QY 330 ----- 329  
DB 376 FLPEFDDPASPNTAPLQPEOLQVETLEETITGYLISAMPDLPDLVSFONLQVIRGRI 435

QY 330 ----- 329  
DB 436 LHNCAVSLTLCGIGISWIGLSRLBSGLALIHNNHLCFVHTVPMQDLFENPHQALLH 495  
QY 330 ----- 329  
DB 496 TANREDECVGBGLACHQLCANGHCWGPPTQCVNCSQFLRGQCEVEECRVLGLPREV 555  
QY 330 ----- 329  
DB 556 NARHCLPCHPECCQPNQSVTCFGEADOCVACAHYKDPFCVAPCPSGVKPDLSYPIWK 615  
QY 330 ----- 329  
DB 616 PPDEGACQPCPINCTHSCVDLDDKCAEQRASPLTISAVGILLVVLGVVFGILI 675  
QY 330 ----- 329  
DB 676 KRQCKIRKRYTRRLQETELVEPLTPSGAMPNQAMRLKETELRKVKVLSGAFGTIVY 735  
QY 330 ----- 329  
DB 736 KGIWIPDEENKIPVAIKVIRENTSPKANKETLDEAVYVAGVSPVSRLLGICLTSTVQ 795  
QY 330 ----- 329  
DB 796 LVTQLMPYGCLLDHVRENKRLGSDLLNMCQIAKMSYLEDVRLVHRDLAARVLYKS 855  
QY 330 ----- 329  
DB 856 PNHVKITDFGLARLLDIDETEHADGKVPKIMMALESILRRFTHQSDVMSYGVYVNL 915  
QY 330 ----- 329  
DB 916 MTEGAKPYDGIIPAREIPDLLEKGERLPPRPICTIDVYIMVCMNIDSECRPRELVE 975  
QY 330 ----- 329  
DB 976 FSRMARDPQRFVYIQNEDLPASPLDSTFYRSLEDDMDGLVDAEYLVYPOGFFCPR 1035  
QY 330 --GAGCMVHHRRSSSTNSGGGDLTLGLEPSEEARPSRLARBSAGSDVDFDGLMGAA 387  
DB 1036 APGAGCMVHHRRSSSTNSGGGDLTLGLEPSEEARPSRLARBSAGSDVDFDGLMGAA 1095  
QY 388 KGLOSLPTNHPSPLORYSEDDPTVPLPSETDGVAPLTCSPQPEYVNOQDPVPRQPSPRBG 447  
DB 1096 KGLOSLPTNHPSPLORYSEDDPTVPLPSETDGVAPLTCSPQPEYVNOQDPVPRQPSPRBG 1155  
QY 448 PLPAARPAAGATLERAKTLPCKNGVYKDFAFAGAVENPEYLTPOGAAPQPHPPAPSP 507  
DB 1156 PLPAARPAAGATLERAKTLPCKNGVYKDFAFAGAVENPEYLTPOGAAPQPHPPAPSP 1215  
QY 508 AFDNLVYWDQDPPERGAPPSFTFGPTAENPEYLCGLDVP 546  
DB 1216 AFDNLVYWDQDPPERGAPPSFTFGPTAENPEYLCGLDVP 1254  
RESULT 12  
AAM92406  
ID AAM92406 standard; Protein: 1255 AA.  
XX AAM92406;  
XX AC  
XX AC  
XX DT 21-Apr-1999 (first entry)  
XX DE Human HER-2/neu oncogene protein.  
XX DE  
XX KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;  
XX KW malignancy; treatment; tumour.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers

FT Region 676..1255  
 XX /note= "region which elicits immune response"  
 PN US5869445-A.  
 XX 09-FEB-1999.  
 PD  
 XX 01-APR-1996; 9605-0625101.  
 PF  
 XX 01-APR-1996; 9605-0625101.  
 PR 01-APR-1996; 9605-0625101.  
 PR 17-MAR-1993; 9305-0033644.  
 PR 12-AUG-1993; 9305-0106112.  
 PR 31-MAR-1995; 9505-0414417.  
 XX  
 PA (UNIV ) UNIV WASHINGTON.  
 PI Cheever MA, Dists ML;  
 XX  
 XX WPI; 1999-152835/13.  
 DR N-PSDB; AAX01912.  
 XX  
 XX Use of HER-2/neu polypeptides - for eliciting an immune response to  
 PT an HER-2/neu associated malignancy, particularly for treating or  
 PT preventing tumours  
 XX  
 XX Claim 3; Column 31-38; 26pp; English.  
 PS  
 XX  
 CC This sequence represents the human HER-2/neu oncogene protein. A fragment  
 CC of this protein is used in a method for eliciting or enhancing an immune  
 CC response to HER-2/neu protein. The polypeptide can stimulate T cells and  
 CC B cells to produce an immune response to the HER-2/neu protein. The  
 CC method can be used for immunisation against a malignancy in which the  
 CC HER-2/neu oncogene is associated and in the treatment of an existing  
 CC tumour, or to prevent tumour occurrence or reoccurrence.  
 CC  
 XX  
 SO Sequence 1255 AA:  
 Query Match 64.38; Score 2405.5; DB 20; Length 1255;  
 Best Local Similarity 41.18; Pred. No. 3.5e-142;  
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 330 ----- 329  
 Db 496 TANPDECEVGEGLACHQLCARGHCWGPPTQVCNCSQFLRGQCEVCECRYLQGLPREYV 555  
 QY 330 ----- 329  
 Db 556 NARHCLFCHPECQPNQSYNCFGEADQCAAHYKDPPECVARGSPGVKPDLSYMPWK 615  
 QY 330 ----- 329  
 Db 616 FPDEGACQPCPINCSTHSCVDLDKCGCPAQBRASPLTISAVGILLVVLGVVFGIL 675  
 QY 330 ----- 329  
 Db 676 KRRQOKIRKTYMRRLQETELVEPLTPSGAMPNQAOKRIKETELRRVYKVGSGAGFTVY 735  
 QY 330 ----- 329  
 Db 736 KGIVIPDGENVKIVAIKVLRENTSPKANKETIDEAVYMGVGSPIYSRLIGICTSTVQ 795  
 QY 330 ----- 329  
 Db 796 LVYQIMPYGCLLDHVRENRGRSGSDDLNCMQIAKGSYLEDVRLVHRDLARNLVKS 855  
 QY 330 ----- 329  
 Db 856 PNHYKITDFGLARLLDIDETEHADGKVPKIMNALESILRRPETHOSDVNSYGVWEL 915  
 QY 330 ----- 329  
 Db 916 MTFCAKPYDGIPIAREIPDLLEKGRLPQPICTIDVYIMINKCMIDSECRPRELVE 975  
 QY 330 ----- 329  
 Db 976 FSRMARDPQRFVYIIONEDLGPASPLDSTFYRSLLLEDDMDGLVDAEELYVQGFPCDP 1035  
 QY 330 ----- 387  
 Db 1036 APGAGGVNHHNRSSSTRSGGDLTLGLEPSEEARPSPLAPSGASDVDDGGLGAA 1095  
 QY 388 KGLSLPTHPSPQKRSSEDPVPLPSETDGVYAPLTCSPQREYVNOPDVAPPPSPREG 447  
 Db 1096 KGLSLPTHPSPQKRSSEDPVPLPSETDGVYAPLTCSPQREYVNOPDVAPPPSPREG 1155  
 QY 448 PLPAARAGATLERAKTLPCKNGVVKDVFAGAVENPEYLTPOGGAAPQHPPPAFSP 507  
 Db 1156 PLPAARAGATLERAKTLPCKNGVVKDVFAGAVENPEYLTPOGGAAPQHPPPAFSP 1215  
 QY 508 AFQNLVYWDODPPERGAAPSTFKGTPTAENPEYLGIDVP 546  
 Db 1216 AFQNLVYWDODPPERGAAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 13  
 ID AAB21198 standard; protein; 1255 AA.  
 XX AAB21198;  
 AC  
 XX  
 XX 12-JAN-2001 (first entry)  
 DT  
 XX  
 DE Human HER-2/neu protein.  
 XX  
 XX Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;  
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;  
 KW colon cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200044899-A1.  
 XX  
 XX 03-AUG-2000.  
 PD  
 XX



PF 28-JAN-2000; 2000MO-US02164.  
 XX  
 PR 29-JAN-1999; 99US-0117976.  
 XX  
 PA (CORI-) CORIXA CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM.  
 XX  
 PI Cheever MA, Gheysen D;  
 XX  
 DR WPI: 2000-505976/45.  
 DR N-PSDB: AAA89736.  
 XX  
 PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins  
 PT useful for vaccinating against breast, ovarian, colon, lung and  
 PT prostate cancers -  
 PS  
 PS Claim 52; Fig 7; 128pp: English.  
 XX  
 CC The present sequence is the human HER-2/neu protein. It is a member of  
 CC the tyrosine kinase family of receptor-like glycoproteins and shows  
 CC homology to the epidermal growth factor receptor (EGFR). It probably  
 CC plays a part in cell growth and/or differentiation. The HER-2/neu  
 CC gene is an oncogene. An HER-2/neu fusion protein comprising a  
 CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation  
 CC domain may be used to treat or prevent cancer by eliciting or  
 CC enhancing an immune response to the HER-2/neu protein. It may be used  
 CC to treat malignancies such as breast, ovarian, colon, lung and  
 CC prostate cancers, and may be used as an antigen to vaccinate against  
 CC these neoplasias.  
 CC  
 XX  
 SQ Sequence 1255 AA:  
 Query Match 64.3%; Score 2405.5; DB 21; Length 1255;  
 Best Local Similarity 41.1%; Pred. No. 3.5e-142;  
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

Db 556 NARHCLPCHPECQPNQSVTCGPEADQVACAHYKIDPPFCVAPRPSGVKPDLSYMPIWK 615  
 QY 330 ----- 329  
 Db 616 FPDEGACQPCRPINCTHSCVDLDKGPABORASPLTISIAYVIGLLVVLGVVFGILI 675  
 QY 330 ----- 329  
 Db 676 KRROOKIKKTYMRRLQETELVEPLTPSGAMPNOAQRILKETELRKVKYLGSQAFGVY 735  
 QY 330 ----- 329  
 Db 736 KGIWIDGENVKIPVAIKVIRENTSPKANKETLDEAYVMAGVSPYVRLGICLTSTVQ 795  
 QY 330 ----- 329  
 Db 796 LVTLMPYGLLDHYRENHRLGSDLLNMCQIAKMSLYLEDVRLVHRDLAARVLYKS 855  
 QY 330 ----- 329  
 Db 856 PNHVKITDFGLARLLDIDETEHADGKVPYIKMALESILRRRTHQSDVMSYCVTVNEL 915  
 QY 330 ----- 329  
 Db 916 MTFGAKPYDGIPIREIPDLLEKGERLPQPICTIDVYIMVNCWMIDSECRPRFRELVS 975  
 QY 330 ----- 329  
 Db 976 FSRMARDPQRFVYIIONEDLGAPSLDSTFYRSLLEDMDGDLVAEEYLVPOQGFCCPD 1035  
 QY 330 --GAGWVHNHRSSSTRSGGDLTLGLEPSEEARSPPLARSBAGSDVFDGDLGMAA 387  
 Db 1036 APGAGWVHNHRSSSTRSGGDLTLGLEPSEEARSPPLARSBAGSDVFDGDLGMAA 1095  
 QY 388 KGLSLPTNDSPLQRYSEDPVPLPSETDGYVAPLTCSPQREYVNPQDVAPRPPREG 447  
 Db 1096 KGLSLPTNDSPLQRYSEDPVPLPSETDGYVAPLTCSPQREYVNPQDVAPRPPREG 1155  
 QY 448 PLPAARPGATLERAKTLSPKNGVYKVFAPGAVENPEYLTPOGGAAPRHPARFSP 507  
 Db 1156 PLPAARPGATLERAKTLSPKNGVYKVFAPGAVENPEYLTPOGGAAPRHPARFSP 1215  
 QY 508 AFDNLVYWDODPPERGAAPPSTFKGTPTAENPEYLGIDVP 546  
 Db 1216 AFDNLVYWDODPPERGAAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 14  
 AAY84780  
 ID AAY84780 standard; Protein: 1255 AA.  
 XX  
 AC AAY84780;  
 XX  
 DT 08-AUG-2000 (first entry)  
 XX  
 DE Amino acid sequence of the SPLICE erbb-2 receptor protein.  
 XX  
 KW SPLICE erbb-2 receptor protein; cell transformation disorder; cancer;  
 KW tumor cell proliferation; tissue degeneration; arthropathy;  
 KW bone resorption; inflammatory disease; degenerative disorder;  
 KW wound healing.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200020579-A1.  
 XX  
 PD 13-Apr-2000.  
 XX  
 PE 01-Oct-1999; 99WO-CA00912.  
 XX  
 PR 02-Oct-1998; 98US-0165192.

PA (UYMC-) UNIT MCMMASTER.  
 XX  
 PI Muller WJ, Siegel PM;  
 XX  
 DR WPI: 2000-303768/26.  
 DR N-PSDB: AAA14812.  
 XX  
 PT Nucleic acid encoding an erbB-2 receptor protein designated SPLICE  
 PT erbB-2, inhibitors of the protein are useful for treatment of cancer  
 XX  
 PS Claim 3: Fig 2: 60pp; English.  
 CC The present sequence represents a SPLICE erbB-2 receptor protein. The  
 CC protein has an in-frame deletion of 16 amino acids, 2 of which are  
 CC conserved cysteine residues, compared to the unspliced protein. The  
 CC erbB-2 polynucleotide is used to construct probes for detecting  
 CC disorders of cell transformation such as cancer. Antibodies to the  
 CC protein may be used to detect SPLICE erbB-2 in a sample. Agents of  
 CC (e.g. antisense oligonucleotides) which inhibit the expression of  
 CC SPLICE erbB-2 are useful for reducing tumor cell proliferation and  
 CC treating cancer. Substances which stimulate SPLICE erbB-2 are useful  
 CC for treating conditions involving damaged cells including conditions  
 CC in which degeneration of tissue occurs, such as arthropathy, bone  
 CC resorption, inflammatory diseases, degenerative disorders of the  
 CC central nervous system and wound healing.  
 XX  
 XX Sequence 1255 AA;

Query Match 64.3%; Score 2405.5; DB 21; Length 1255;  
 Best Local Similarity 41.1%; Pred. No. 3,5e-142;  
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGASTQVCTGDMKRLRPASPEHLDMLRLHYGCGVVOGNLELTYLPTNASLSFLQ 94  
 DB 16 LPGAASSTQVCTGDMKRLRPASPEHLDMLRLHYGCGVVOGNLELTYLPTNASLSFLQ 75  
 QY 95 DICEVQSVYLIANOVQVQLRLRYRGTOLEPDNALVLDNGDPLNNTPTTGASPG 154  
 DB 76 DICEVQSVYLIANOVQVQLRLRYRGTOLEPDNALVLDNGDPLNNTPTTGASPG 135  
 QY 155 GLRELQRLSTLEILKGGVLIQIRNPQLCYQDTILMKDIFHKNQALATLIDTNRRAACHPC 214  
 DB 136 GLRELQRLSTLEILKGGVLIQIRNPQLCYQDTILMKDIFHKNQALATLIDTNRRAACHPC 195  
 QY 215 SPCKSGRCGESESDCQSLRTVYACGACGACRCKPLPTDCHEQCAAGCTGPKHSDCLAC 274  
 DB 196 SPCKSGRCGESESDCQSLRTVYACGACGACRCKPLPTDCHEQCAAGCTGPKHSDCLAC 255  
 QY 275 LHFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYVYLSYDVS----- 329  
 DB 256 LHFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYVYLSYDVSCTLYVC 315  
 QY 330 ----- 329  
 DB 316 PLHNQEVTAEDGTGRCEKSKPCARVCYGLMEHLREVRVAVTSANIOEFAGCKKIFESLA 375  
 QY 330 ----- 329  
 DB 376 FLPESEFDGDPASNTAPLQPEQLQVETLEETIGLYISAMPDSLPLDSVFONQIVIRGI 435  
 QY 330 ----- 329  
 DB 436 LHNQANSSTLQGLISWLGRLSRELGSLALIHNTHTLCFVHTVPWDQLFRNHQALLH 495  
 QY 330 ----- 329  
 DB 496 TANRPEDECVGEGLAGHQLCARGHGNGRPTQVCNCSQFLRGQECVECVLQGLPREYV 555  
 QY 330 ----- 329  
 DB 556 NARHCLPCHPECOFONGSVTCGFEADQCAHAHYKDPFCVACRPSGVKPDLSYPMWK 615  
 QY 330 ----- 329

DB 616 FPDEGACQPCPINCSTHSCVDLDDKGCAPQERASPLTSIISAVVIGILLVVLGVFGLI 675  
 QY 330 ----- 329  
 DB 676 KRROQKIRKTYMRRLAETELVEPLTPSGAMPNOAQRILKETELRKVKVLSGAGCTGY 735  
 QY 330 ----- 329  
 DB 736 KGIWIPGENVKIPVAIKVLRENTSPKANKEILDEAVVMAGVGSPPVSRLLGLCTSTVQ 795  
 QY 330 ----- 329  
 DB 796 LVTQLMRYGCLLDHYRENRGLGSQDLLMCMQIAGKMSYLEDVRLVHRDLAARNVLKS 855  
 QY 330 ----- 329  
 DB 856 PNHVKTIDFGARLIDIDETEHADGKVPKMMALESILRRRTHOSDWSYGVYWEI 915  
 QY 330 ----- 329  
 DB 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVYKCMWIDSECPREFRELYSE 975  
 QY 330 ----- 329  
 DB 976 FSRNARDPQREVIQNEDELGPASPLDSTFYRSLLEDMDGLVDAEELYVPQGFCCPDP 1035  
 QY 330 --GAGGVNHRHRSSSTRSGGDLTLGLEPSEEPAPRSLAPSEAGSDVFDGLGNGAA 387  
 DB 1036 APGAGVNHHRHRSSSTRSGGDLTLGLEPSEEPAPRSLAPSEAGSDVFDGLGNGAA 1095  
 QY 388 KGLOSPLTHDPSLQRYSEDPVPLPSETDGYVAPLTCSPQPEVYNQPDVRRPQPSREG 447  
 DB 1096 KGLOSPLTHDPSLQRYSEDPVPLPSETDGYVAPLTCSPQPEVYNQPDVRRPQPSREG 1155  
 QY 448 PLPARPAGATLERAKTLSPGKNGVYKDVAFAGAVENPEYLTPQGAAPQHPPPAPSP 507  
 DB 1156 PLPARPAGATLERAKTLSPGKNGVYKDVAFAGAVENPEYLTPQGAAPQHPPPAPSP 1215  
 QY 508 AFDNLVYWDQDPPERGAPSPSTFKGTPPAENPEYGLDVP 546  
 DB 1216 AFDNLVYWDQDPPERGAPSPSTFKGTPPAENPEYGLDVP 1254

RESULT 15  
 AAB85458  
 ID AAB85458 standard; Protein: 1255 AA.  
 XX  
 AC AAB85458;  
 XX  
 DT 25-SEP-2001 (first entry)  
 XX  
 DE Human HER-2/neu protein.  
 XX  
 KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu;  
 KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153463-A2.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 19-JAN-2001; 2001WO-US01850.  
 XX  
 PR 21-JAN-2000; 2000US-0177545.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Cheever MA, Hand-Zimmermann S;  
 XX  
 DR WPI: 2001-476112/51.  
 DR N-PSDB: AAB23392.

XX New antigen-presenting cells, useful as vaccines for eliciting or  
 PT enhancing an immune response to HER-2/neu protein, particularly useful  
 PT for treating or preventing cancer, e.g. breast cancer  
 XX  
 PS Claim 2; Page 41-46; 49pp; English.  
 XX  
 CC The invention provides an isolated antigen-presenting cell, which  
 CC expresses at least an immunogenic portion of a polypeptide that produces  
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are  
 CC useful as vaccines for eliciting or enhancing an immune response to  
 CC HER-2/neu protein, particularly in treating or preventing malignancies in  
 CC which the HER-2/neu oncogene is associated. Specifically, these are  
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,  
 CC colon, lung or prostate cancers. The present sequence represents  
 CC the human HER-2/neu protein (also known as p185 or c-erbB2).  
 CC  
 XX  
 SQ Sequence 1255 AA;  
 Query Match 64.3%; Score 2405.5; DB 22; Length 1255;  
 Best Local Similarity 41.1%; Pred. No. 3.5e-142;  
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;  
 QY 35 LARGASTQVCTGDMKRLPASPEHLMRLHLYGCGVVOGNELEYLPTNASLSFLQ 94  
 DB 16 LPPGASTQVCTGDMKRLPASPEHLMRLHLYGCGVVOGNELEYLPTNASLSFLQ 75  
 QY 95 DIOEVGVLLAHNOVROVRLRIYRGTOLEEDNVALLVNDNDPLNTTPVYGASPG 154  
 DB 76 DIOEVGVLLAHNOVROVRLRIYRGTOLEEDNVALLVNDNDPLNTTPVYGASPG 135  
 QY 155 GLRELQLSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQLATLIDTNRSHACHP 214  
 DB 136 GLRELQLSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQLATLIDTNRSHACHP 195  
 QY 215 SPMCKSGKMGESSEDCOSTRTVCAGGACARCKPLPDCCHEGCAAGCTGPKHSDCIAC 274  
 DB 196 SPMCKSGKMGESSEDCOSTRTVCAGGACARCKPLPDCCHEGCAAGCTGPKHSDCIAC 255  
 QY 275 LHFNSGICELHCPALVYNTDTPESMPNPGRYTFEGASCTACPYNTLSDVGS ---- 329  
 DB 256 LHFNSGICELHCPALVYNTDTPESMPNPGRYTFEGASCTACPYNTLSDVGSCTLV 315  
 QY 330 ----- 329  
 DB 316 PLHNGEYTAEDGTQCEKSKPCARVCYGLGMEHLREVRATVANSIOEFAGCKKIFGSLA 375  
 QY 330 ----- 329  
 DB 376 FLPEFDDPASNTAPLQPEQLQVEETLEITGYLYISAMPDLSPLDSVFQNLQVIRGRI 435  
 QY 330 ----- 329  
 DB 436 LHNQAYSLTLOGLISWLGSLRELGSGLAIHHNTHLCFVHTVPWDQLFRNPQALLH 495  
 QY 330 ----- 329  
 DB 496 TANREDCEVGEGLACHOLCARGHCMGPGPTQCVNCSQPLRGQCEVEECRYLOGIIPREYV 555  
 QY 330 ----- 329  
 DB 556 NARHCLPCHPCQPNQSVTCFGEADQVACAHYKDPFCVARGPSGVKPDLSYPIWK 615  
 QY 330 ----- 329  
 DB 616 FPDEGACQPCPINCTHSCVDLDKCPAQASPLTSIIISAVVIGLLVVLGVVFGILI 675  
 QY 330 ----- 329  
 DB 676 KRROOKIKKTYMRRLQETELVEPLTPSGAPNQAQRIKLETLRKVKVVLGSGAFGIVY 735  
 QY 330 ----- 329

DB 736 KGIWIPDGENVKIPVAIKVIRENTSPKANKIELDEAYVMAGVSPVYSRLGICITSTVQ 795  
 QY 330 ----- 329  
 DB 796 LVTOIMPYGCLLDHVRENKRLGSODLLNMCQIAKMSYLEDVRLVHRDLAARNVLVKS 855  
 QY 330 ----- 329  
 DB 856 PNHYKITDFGLARLLIDIDETEHADGCKVPIKMMALLESILARRRPTHOSDVMSTVWEL 915  
 QY 330 ----- 329  
 DB 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVNCWMIDSECRPRRELIVSE 975  
 QY 330 ----- 329  
 DB 976 FSRMARDPQREVVIONEDLGASPLDSTFTFRSLLEDMDGDLVDAEYLVPOGFFCPDP 1035  
 QY 330 --GAGGVNHRHRRSSSTRSGGDLTLGLEPSEEEAPRSPLAPSEAGSDVFDGLMGAA 387  
 DB 1036 APGAGGVNHRHRRSSSTRSGGDLTLGLEPSEEEAPRSPLAPSEAGSDVFDGLMGAA 1095  
 QY 388 KGLQSLPTHDSPLQRTSEDTVPPLPSETDGVAPLTCSPQREYVNOVDVAPRPPSPREG 447  
 DB 1096 KGLQSLPTHDSPLQRTSEDTVPPLPSETDGVAPLTCSPQREYVNOVDVAPRPPSPREG 1155  
 QY 448 PLPAARPAATLEAKTLSPGKNGVVDVAFAGAVENPEYLTPOGGAAPRPPPAFSP 507  
 DB 1156 PLPAARPAATLEAKTLSPGKNGVVDVAFAGAVENPEYLTPOGGAAPRPPPAFSP 1215  
 QY 508 AFDNLYWDDPPPERGAPSTFKGTPTAENPEYLGIDVP 546  
 DB 1216 AFDNLYWDDPPPERGAPSTFKGTPTAENPEYLGIDVP 1254

Search completed: April 28, 2003, 13:39:43  
 Job time : 72.3846 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:38:09 ; Search time 7.98328 Seconds

(Without alignments)  
2613.108 Million cell updates/sec

Title: US-09-821-883-25

Sequence: 1 GAGGWHHRHRRSSSTRSGG.....STFKGPTAENPEYLGLDVP 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR\_73:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1177	99.6	1255	1 A24571	protein-tyrosine k
2	966.5	81.8	1254	2 A48161	p-185 precursor
3	954	80.7	1260	1 TVR1ND	protein-tyrosine k
4	164	13.9	1210	1 GQHUE	epidermal growth f
5	157	13.3	1308	2 A47253	epidermal growth f
6	145.5	12.3	707	2 A46302	PTB-associated spl
7	138	11.7	698	1 TVEFVY	protein-tyrosine k
8	138	11.7	1223	1 TYCHIV	epidermal growth f
9	137.5	11.6	392	1 P1H06	salivary proline-r
10	135.5	11.5	553	1 H70766	probable dihydroli
11	135	11.4	620	2 S06733	hydroxyproline-ric
12	134.5	11.4	379	2 S31719	proline-rich prote
13	134	11.3	309	2 T29293	hypothetical prote
14	134	11.3	674	2 T05264	probable serine/th
15	134	11.3	1210	2 A53183	epidermal growth f
16	132.5	11.2	442	2 S25299	cell wall glycopro
17	132.5	11.2	442	2 S50062	extensin precursor
18	131.5	11.1	473	2 S50755	hypothetical prote
19	130	11.0	847	2 P96531	hypothetical prote
20	129.5	11.0	251	1 P1H06	salivary proline-r
21	129.5	11.0	476	2 T27051	hypothetical prote
22	129	10.9	440	2 JC7807	Wiskott-Aldrich sy
23	127.5	10.8	403	2 S52796	prp12 protein - hu
24	127.5	10.8	3149	1 O0B8	BPLF1 protein - hu
25	127	10.7	369	2 S20500	hydroxyproline-ric
26	126	10.7	990	2 T14756	hypothetical prote
27	126	10.7	4957	2 T03455	ALK protein - huma
28	126	10.7	5262	2 T03454	ALK protein - huma
29	125.5	10.6	760	2 T06291	extensin homolog T

30	125	10.6	1460	1 EDBE1F	Immediate-early pr
31	124.5	10.5	775	1 EDBE11	Immediate-early pr
32	124	10.5	551	2 S57447	HPBR1-7 protein - r
33	124	10.5	1106	1 TVH0CL	transforming prote
34	123.5	10.4	1494	2 T14355	protein-tyrosine-p
35	123	10.4	2715	2 T13049	eyellid - fruit fly
36	122	10.3	300	2 A41870	dna protein - Str
37	121.5	10.3	368	2 C29356	hydroxyproline-ric
38	121.5	10.3	368	2 C29356	hydroxyproline-ric
39	121.5	10.3	938	1 Q0B824	nuclear antigen EB
40	121.5	10.3	1357	2 T29265	hypothetical prote
41	121	10.2	1585	2 T31611	hypothetical prote
42	120.5	10.2	301	2 T21311	hypothetical prote
43	120.5	10.2	424	2 T33663	hypothetical prote
44	120.5	10.2	449	2 S16748	proline-rich prote
45	120	10.2	318	2 T29479	hypothetical prote

## ALIGNMENTS

RESULT 1  
A24571  
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human  
N:Alternate names: c-erb-B-2 protein precursor; Kinase-related transforming protein e  
C:Species: Homo sapiens (man)  
C:Date: 25-Oct-1987 #sequence:revision 06-Dec-1996 #text:change 11-Jun-1999  
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622  
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Salto, T  
Nature 319, 230-234, 1986  
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth.  
A:Reference number: A24571; MUID:86118663; PMID:3003577  
A:Accession: A24571  
A:Molecule type: mRNA  
A:Residues: 1-1255 <YAM>  
A:Cross-references: GB:X03363; NID:931197; PIDN:CAA27060.1; PID:931198  
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985  
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid  
A:Reference number: A25491; MUID:86016729; PMID:2995967  
A:Accession: A25491  
A:Molecule type: DNA  
A:Residues: 737-1031 <SEM>  
A:Cross-references: GB:M11767; NID:9182163; PIDN:AAA35808.1; PID:9553282  
R:Cozens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg  
Science 230, 1132-1139, 1985  
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro  
A:Accession: A44188  
A:Reference number: A44188; MUID:86070181; PMID:2999974  
A:Molecule type: DNA  
A:Residues: 740-910 <COU1>  
A:Cross-references: GB:M12036; NID:9183988; PIDN:AAA35978.1; PID:9183989  
A:Accession: B44188  
A:Molecule type: mRNA  
A:Residues: 1-517, 'RALM', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>  
A:Cross-references: GB:M11730; NID:9183986  
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.  
Science 229, 974-976, 1985  
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.  
A:Reference number: I59509; MUID:85272597; PMID:2992089  
A:Accession: I59509  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 832-909 <REX>  
A:Cross-references: GB:I29395; NID:9459807; PIDN:AAA35809.1; PID:9459808  
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.  
Mol. Cell. Biol. 7, 2597-2601, 1987  
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio  
A:Reference number: I57622; MUID:87286898; PMID:3039351  
A:Accession: I57622  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:q183983; PIDN:AAA58637.1; PID:9553332  
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30  
 C:Genetics:  
 A:Gene: GDB:ERB2; NGL: NEU; HER-2  
 A:Cross-references: GDB:120613; OMIM:164870  
 A:Map position: 17q21.1-17q21.1  
 A:Introns: 25/1; 75/3; 147/1; 883/3  
 A>Note: the list of introns is incomplete  
 C:Function:  
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-125/Domain: product: protein-tyrosine kinase erbB2 #status predicted <MAT>  
 F:22-653/Domain: extracellular #status predicted <EXT>  
 F:70-304/Domain: EGF receptor extracellular domain repeat <EEL>  
 F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>  
 F:554-675/Domain: transmembrane #status predicted <TM>  
 F:576-1235/Domain: intracellular #status predicted <INT>  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif  
 F:68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:86/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F:753/Active site: lys #status predicted  
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 99.6%; Score 1177; DB 1; Length 1255;  
 Best Local Similarity 99.5%; Pred. No. 1,4e-69;  
 Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPRSLAPSEAGSDVFDGDLGMAAKG 60  
 DB 1038 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPRSLAPSEAGSDVFDGDLGMAAKG 1097

QY 61 LOSLPHDPSPLQRYSDPPLPSEEDGVAVPLTCSPOPEYVNOPDVRQPPSPREGPL 120  
 DB 1098 LOSLPHDPSPLQRYSDPPLPSEEDGVAVPLTCSPOPEYVNOPDVRQPPSPREGPL 1157

QY 121 PAARPGATLERAKTISPGKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAPF 180  
 DB 1158 PAARPGATLERAKTISPGKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAPF 1217

QY 181 DNLTYMDQDPERGAPSPSTFKGTPTAENPEYLTGLDVP 217  
 DB 1218 DNLTYMDQDPERGAPSPSTFKGTPTAENPEYLTGLDVP 1254

RESULT 2  
 I48161  
 P-185 precursor - golden hamster  
 C:Species: Mesocricetus auratus (golden hamster)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Jun-1999  
 C:Accession: I48161  
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa, Gene 140, 251-255, 1999  
 A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.  
 A:Reference number: I48161; MUID:94193007; PMID:7908275  
 A:Accession: I48161  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1254 <RES>  
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595  
 C:Genetics:  
 A:Gene: neu  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP  
 F:718-983/Domain: protein kinase homology <KIN>  
 F:726-734/Region: protein kinase ATP-binding motif  
 Query Match 81.8%; Score 966.5; DB 2; Length 1254;  
 Best Local Similarity 82.0%; Pred. No. 8.1e-56;  
 Matches 178; Conservative 13; Mismatches 25; Indels 1; Gaps 1;

QY 1 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPRSLAPSEAGSDVFDGDLGMAAKG 60  
 DB 1038 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPRSLAPSEAGSDVFDGDLGMAAKG 1097

QY 61 LOSLPHDPSPLQRYSDPPLPSEEDGVAVPLTCSPOPEYVNOPDVRQPPSPREGPL 120  
 DB 1098 LOSLPHDPSPLQRYSDPPLPSEEDGVAVPLTCSPOPEYVNOPDVRQPPSPREGPL 1157

QY 121 PAARPGATLERAKTISPGKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAPF 180  
 DB 1158 PAARPGATLERAKTISPGKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAPF 1216

QY 181 DNLTYMDQDPERGAPSPSTFKGTPTAENPEYLTGLDVP 217  
 DB 1217 DNLTYMDQDPERGAPSPSTFKGTPTAENPEYLTGLDVP 1253

RESULT 3  
 TVERNU  
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 11-Jun-1999  
 C:Accession: A24562; A61204  
 R:Barigmann, C.I.; Hung, M.C.; Weinberg, R.A.  
 Nature 319, 226-230, 1986  
 A:Title: The new oncogene encodes an epidermal growth factor receptor-related protein  
 A:Reference number: A24562; MUID:86118662; PMID:3945311  
 A:Accession: A24562  
 A:Molecule type: mRNA  
 A:Residues: 1-1260 <BAR>  
 A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746  
 R:Masu, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen Carcinogenesis 12, 1975-1978, 1991  
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals a 2-chitzolyllformamide or N-methyl-N-nitrosocoumarin  
 A:Reference number: A61204; MUID:92035293; PMID:1682063  
 A:Accession: A61204  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 637-663, 'V', 665-702 <MAS>  
 A>Note: authors translated the codon GCA for residue 25 as Val  
 C:Genetics:  
 A:Gene: neu  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase  
 F:1-19/Domain: signal sequence #status predicted <SIG>  
 F:20-1260/Domain: product: protein-tyrosine kinase neu #status predicted <MAT>  
 F:558-680/Domain: transmembrane #status predicted <TM>  
 F:723-988/Domain: protein kinase homology <KIN>  
 F:731-739/Region: protein kinase ATP-binding motif  
 F:71,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted  
 F:758/Active site: lys #status predicted  
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 80.7%; Score 954; DB 1; Length 1260;  
 Best Local Similarity 81.6%; Pred. No. 5.4e-55;  
 Matches 177; Conservative 8; Mismatches 32; Indels 0; Gaps 0;

QY 1 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPRSLAPSEAGSDVFDGDLGMAAKG 60  
 DB 1043 GTSTAHRRSSSTRSGGDLTLGLEPSEEARPRSLAPSEAGSDVFDGDLGMAAKG 1102

QY 61 LOSLPHDPSPLQRYSDPPLPSEEDGVAVPLTCSPOPEYVNOPDVRQPPSPREGPL 120  
 DB 1103 LOSLPHDPSPLQRYSDPPLPSEEDGVAVPLTCSPOPEYVNOPDVRQPPSPREGPL 1162

QY 121 PAARPGATLERAKTISPGKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAPF 180  
 DB 1163 PAARPGATLERAKTISPGKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAPF 1222

QY 181 DNLTYMDQDPERGAPSPSTFKGTPTAENPEYLTGLDVP 217

Db 1223 DMLTYMDNSSEOGPPSNFEGTPTAENPEYLGDVP 1259

# RESULT 4

epidermal growth factor receptor precursor - human

N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erDB

C:Species: Homo sapiens (man)

C:Date: 15-Nov-1984 #sequence\_revision 27-Nov-1985 #text\_change 11-Jun-1999

C:Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05281; A60143; A3

R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y

rg, P.H.

Nature 309, 418-425, 1984

A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of

A:Accession number: A00641; MUID:84219729; PMID:6328312

A:Accession: A00641

A:Molecule type: mRNA

A:Residues: 1-1210 <L>

A:Cross-references: EMBL:X00588; NID:g31113; PIDN:CAA25240.1; PID:g57924

A:Note: the authors translated the codon AAG for residue 540 as Asn

R:Jihli, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.

Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985

A:Title: Characterization and sequence of the promoter region of the human epidermal gtr

A:Reference number: A25772; MUID:85270438; PMID:2991899

A:Accession: A25772

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-29 <IS>

A:Cross-references: GB:M1134; NID:g31118; PIDN:AAA5370.1; PID:g553272

A:Accession: A05281

A:Molecule type: protein

A:Residues: 25-307; S' 32-51; 454-467 <WEB>

R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.

J. Biol. Chem. 260, 5205-5208, 1985

A:Title: Identification of residues in the nucleotide binding site of the epidermal g

A:Reference number: A60143; MUID:85182650; PMID:2985580

A:Accession: A60143

A:Molecule type: protein

A:Residues: 740-744; X' 746-747 <RUS>

R:Krocowski, B.; Mosig, G.; Cohen, S.

Nature 309, 270-273, 1984

A:Title: ATP-stimulated interaction between epidermal growth factor receptor and supe

A:Reference number: A38023; MUID:84191554; PMID:6325948

A:Contents: annotation; receptor activity

A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA

R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C

Cell 59, 33-43, 1989

A:Title: Functional independence of the epidermal growth factor receptor from a domai

A:Reference number: A33331; MUID:9003233; PMID:2790960

A:Contents: annotation; internalization signal

C:Comments: Binding of EGF to the receptor leads to internalization of the EGF-recepto

C:Genetics:

A:Gene: EGFR

A:Cross-references: GDB:120610; OMIM:131550

A:Map position: 7p12.3-7p12.1

A:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-1210/Product: EGF receptor #status predicted <MAT>

F:25-645/Domain: extracellular #status predicted <EXT>

F:75-300/Domain: EGF receptor extracellular domain repeat <EE1>

F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>

F:646-668/Domain: transmembrane #status predicted <TM>

F:669-1210/Domain: intracellular #status predicted <INT>

F:710-975/Domain: protein kinase homology <KIN>

F:718-726/Region: protein kinase ATP-binding motif

F:999-1046/Region: coated-pit mediated internalization signal

F:1047-1210/Region: inhibitory

F:128,175,352,413,444,528,603/binding site: carbohydrate (Asn) (covalent) #status pre

F:745/Active site: lys #status experimental

Query Match 13.9%; Score 164; DB 1; Length 1210;

Best Local Similarity 30.5%; Pred. No. 0.0026;

Matches 62; Conservative 24; Mismatches 69; Indels 48; Gaps 11;

Qy 29 SEEPAPNSPLAPSGAGSDVFDGLGMAKGLDLPHPSPLOKRSDFVPLPSET- 87

Db 1025 SSPSTSRPLLSLSATSN--NSTVACIDRNGLCSCPKEDSFQRYSSDPTGALTDESI 1082

Qy 88 -DGVVAPLTCSPOPEYVNPQVRRQPPSPRCGLPAPAPGATLEAKTISPGKNGVYKD 146

Db 1083 DDTFL-----PVPEYINQ--SVPRKPGSVONPYHNOPLNP-----APSDPHYQD 1127

Qy 147 VFAAGCAVENEYV--TPOGGAAPQHPHPAPSPAFDNLNLYMDQ-----DP----- 190

Db 1128 --PSTAVAGNEVINTYQ-----PTCVNSFTFSDPAHMAQGSQHSILDNDYQDDE 1176

Qy 191 -PENGAPSPFTKGTPTAENPEYL 212

Db 1177 FPKAKKNGIFKGS-TAENAEYL 1198

RESULT 5

A47253

epidermal growth factor receptor, HER4 - human

C:Species: Homo sapiens (man)

C:Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Jun-1999

C:Accession: A47253

R:Plowman, G.D.; Culouscou, J.M.; Whitney, G.S.; Green, J.M.; Carlson, G.W.; Foy, L.

Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993

A:Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epiderm

A:Reference number: A47253; MUID:93189574; PMID:8383326

A:Accession: A47253  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-1308 <PLOC>  
 A:Cross-references: GB:107868; NID:9337359; PIDN:AA859446.1; PID:9337360  
 A:Note: sequence extracted from NCBI backbone (NCBIP:126842)  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; growth factor receptor  
 F:716-981/Domain: protein kinase homology <KIN>  
 F:724-732/Region: protein kinase ATP-binding motif

Query Match 13.3%; Score 157; DB 2; Length 1308;  
 Best Local Similarity 24.2%; Pred. No. 0.0082;  
 Matches 62; Conservative 23; Mismatches 75; Indels 96; Gaps 10;

QY 16 RSGGCDLTGLLE-PSEEARSPPLAP-SEGAGSDVFDGDMGAKGLQSLPTHPDPSPLQ 73  
 DB 1067 RDGFAEGAGVSVPRAPSTIPEAPVAGATATETDSCNCTLRKRVAPHQEDSDSTQ 1126  
 QY 74 RYSEDPPVPLPS-----ETDGYVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPA 126  
 DB 1127 RYADPPVFAPEKSPRGELDEEGYMTMRDCKPQOXYLMPVE-----ENPFVSR-- 1175  
 QY 127 GATLERAKTLSPGKNGVVKVFAFGAVENPEYLTPOGGAAPQPHPPA----- 175  
 DB 1176 -----KNGDLQ-----ALDNPETHNMSNG-----PKADEYVNEPELYL 1209  
 QY 176 -----FSPAFDNLVYWDODPPERGA--PPTFKGPT----- 205  
 DB 1210 NTFANTLGAKEYLKNILMSPEKAKKAFDNPDIWNHSLPRSTLQHDYDQEXTKTFYK 1269  
 QY 206 -----AENPEYL 212  
 DB 1270 QNGRIRPIVAENPEYL 1285

## RESULT 6

A46302  
 PRT-associated splicing factor, long form - human  
 N:Alternate names: myoblast cell surface antigen 24.1D5; polypyrimidine tract-binding, pC  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Sep-1993 #sequence revision 18-Nov-1994 #text\_change 24-Sep-1999  
 C:Accession: A46302; A43557; S29985  
 R:Patton, J.G.; Porro, E.B.; Galceran, J.; Tempst, P.; Nadal-Ginard, B.  
 Genes Dev. 7, 393-406, 1993  
 A:Title: Cloning and characterization of pSF, a novel pre-mRNA splicing factor.  
 A:Reference number: A46302; MUID:93194059; PMID:8449401  
 A:Accession: A46302  
 A:Molecule type: mRNA  
 A:Residues: 1-707 <PRT>  
 A:Cross-references: EMBL:X70944; NID:938457; PIDN:CAA50283.1; PID:938458  
 A:Note: Sequence extracted from NCBI backbone (NCBIP:127206)  
 R:Gower, H.J.; Moore, S.E.; Dickson, G.; Elsom, V.L.; Nayak, R.; Walsh, F.S.  
 Development 105, 723-731, 1989  
 A:Title: Cloning and characterization of a myoblast cell surface antigen defined by 24.1  
 A:Reference number: A43557; MUID:90091812; PMID:2480877  
 A:Accession: A43557  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 312-707 <GOW>  
 A:Cross-references: GB:X16850; NID:923711; PIDN:CAA34747.1; PID:923712  
 C:Genetics:  
 A:Gene: GDB:SFPO; PSF  
 A:Cross-references: GDB:138275  
 A:Map position: 4q-4q  
 C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein  
 C:Keywords: alternative splicing; pre-mRNA splicing; surface antigen  
 F:298-359/Domain: ribonucleoprotein repeat homology <RRM1>  
 F:372-438/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 12.3%; Score 145.5; DB 2; Length 707;  
 Best Local Similarity 26.6%; Pred. No. 0.024;  
 Matches 58; Conservative 13; Mismatches 82; Indels 65; Gaps 10;

QY 1 GAGGVHHRHRSSTRGGDLTLGLEPSEEARSPPLAPSEGAGSDVFDGDMGAKG 60  
 DB 11 GGGGGGHR-----GGGGGGLHDFR-----SPPGMGLNQNGRPGPGGS 54  
 QY 61 LQSLPTHPDPSPLQRYSE-DPTVPLPSETDGYVAPLTCSPQPEYVNOQDVRPQPPSPRE-- 117  
 DB 55 GPRPPIPPPPHQQOQPPQOQPPQPPHQP--PHQPHQOQPP--PPPODSS 107  
 QY 118 -----GPLPARPAGATLERAKTLSPGKNGVVKVFAFGAVENPEYLTPOGGAAPQ 169  
 DB 108 KPVVAGCGPAPGVGASPPAASSAPRATPPTSA-----PPGSGGPT 150  
 QY 170 PHPPAFSPAFDNLVYWDODPPERGAPST--FKGPT 205  
 DB 151 PRPPPAVTS-----PP--GAPPPPPSSGVPT 176

## RESULT 7

TVEVLV  
 Protein-tyrosine kinase (EC 2.7.1.112) erbB - avian leukosis virus  
 N:Contains: amino end of gag protein; env protein fragment; protein-tyrosine kinase  
 C:Species: avian leukosis virus, ALV  
 C:Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text\_change 11-Jun-1999  
 C:Accession: B00643; A00643  
 R:Nilsson, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Ralnes  
 Cell 41, 719-726, 1985  
 A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and  
 A:Reference number: A00643; MUID:85228222; PMID:2988784  
 A:Accession: B00643  
 A:Molecule type: mRNA  
 A:Residues: 1-698 <NIL>  
 A:Cross-references: GB:M10066; GB:M13881; NID:9211749; PIDN:AAA48763.1; PID:9211750  
 A:Note: In Genbank entry CHKERBBF, release 109.0, the source is designated as Gallus  
 C:Comment: This protein is synthesized as a gag-env-erbB protein.  
 C:Genetics:  
 A:Gene: gag-env-erbB  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific  
 F:1-6/Product: gag protein (fragment) #status predicted <GAG>  
 F:7-59/Product: env protein (fragment) #status predicted <ENV>  
 F:60-698/Product: protein-tyrosine kinase erbB #status predicted <ERB>  
 F:194-459/Domain: protein kinase homology <KIN>  
 F:202-210/Region: protein kinase ATP-binding motif  
 F:229/Active site: Lys #status predicted

Query Match 11.7%; Score 138; DB 1; Length 698;  
 Best Local Similarity 24.5%; Pred. No. 0.073;  
 Matches 54; Conservative 25; Mismatches 59; Indels 62; Gaps 10;

QY 35 RSPV-----APSEGAGSDVFDGDMGAKGLQSLPTHPDPSPLQRISEDPVPLPSET-- 87  
 DB 515 RTFLSLSLATSNMSATCID-----RNGGHPVREDSFQORSSDPTGNFLESID 566  
 QY 88 DGVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPAATLERAKTLSPGKNGVVKV 147  
 DB 567 DGFL-----PAPEYVNO--LMPKPS-----TAMVNOYNNI 597  
 QY 148 F-----AFGAVENPEYLTPOGGAAPQPHPPAFSPAFDNLVYWDQ----- 188  
 DB 598 SLTAISKLPMDSKYKQNSHTAVDNPEYL-----NTNOSPLAKYFESSPIWIOSGN 649  
 QY 189 ----DPPE-----RGAPPSTFKGPTTAENPEYLGLDVP 217  
 DB 650 QINLNDNDYQDPLPNETKPNGLKTPAAMNPYLRAAP 669

## RESULT 8

TYCHLV  
 Epidermal growth factor receptor precursor - chicken  
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB  
 C:Species: Gallus gallus (chicken)  
 C:Date: 28-Feb-1986 #sequence revision 05-May-1995 #text\_change 04-Feb-2000



A:Experimental source: subject C.J. (large allele)  
A:Accession: CA0750  
A:Molecule type: DNA  
A:Residues: 35-127,'R',129-148,'R',150-151,153-187,'K',189-272,'S',274-336,'S',338-399  
A:Cross-references: GB:S62929  
A:Experimental source: subject M.V.O. (large allele)  
A:Accession: A40750  
A:Molecule type: DNA  
A:Residues: 35-183,245-270,'Q',272-392 <AZ3>  
A:Cross-references: GB:S62928  
A:Experimental source: subject C.J. (medium allele)  
A:Note: authors translated the codon CAA for residue 272 as Arg  
R:Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.  
J. Biol. Chem. 260, 11123-11130, 1985  
A:Title: Differential RNA splicing and post-translational cleavages in the human salivary  
A:Reference number: A92492; MUID:65285325; PMID:2993301  
A:Accession: C25372  
A:Molecule type: mRNA  
A:Residues: 1-183,245-392 <MAE>  
A:Cross-references: GB:K03204; NID:g190485; PIDN:AAA60185.1; PID:g190486  
A:Note: alternatively splice forms lacking portions of the repeat region were also found  
R:Lyons, K.M.; Stein, J.H.; Smithies, O.  
Genetics 120, 267-278, 1988  
A:Title: Length polymorphisms in human proline-rich protein genes generated by Intergenic  
A:Reference number: S02127; MUID:89121440; PMID:2851479  
A:Accession: S02128  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 35-127,250-273,'R',275-277,'R',279-336,'S',338-392 <LYO>  
A:Cross-references: EMBL:X07517  
A:Accession: S02127  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 35-183,245-392 <LY2>  
A:Cross-references: EMBL:X07516  
R:Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P.  
Biochem. J. 235, 2387-2392, 1986  
A:Title: Basic proline-rich proteins from human parotid saliva: complete covalent structure  
A:Reference number: A90502; MUID:86243355; PMID:3521730  
A:Accession: A03293  
A:Molecule type: protein  
A:Residues: 17-38,'AP',41-51,92-148,'R',150-152 <KA2>  
A:Note: among nine basic proline-rich peptides isolated from the saliva, this peptide was  
A:Accession: A90502  
A:Molecule type: protein  
A:Residues: 275-336,'S',338-392 <KAU>  
R:Saitoh, E.; Isemura, S.; Sanada, K.  
J. Biochem. 94, 1991-1999, 1983  
A:Title: Further fractionation of basic proline-rich peptides from human parotid saliva  
A:Reference number: A91974; MUID:84161824; PMID:6671974  
A:Contents: P-H  
A:Accession: A91974  
A:Molecule type: protein  
A:Residues: 'S',338-392 <SAI>  
R:Azen, E.; Lyons, K.M.; McConigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Vann  
Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984  
A:Reference number: A94005; MUID:84298176; PMID:6089212  
A:Accession: A05261  
A:Molecule type: DNA  
A:Residues: 35-39,'P',41-84,'G',86,'R',87-154,'R',218-246;300-306,'T',308-329,'C',331  
A:Accession: A05262  
A:Molecule type: DNA  
A:Residues: 'N',57-59,'A',61-69;334-336,'S',338-339,'R',341-392 <AZ5>  
R:Kauffman, D.; Wong, R.; Bennick, A.; Keller, P.  
Biochemistry 21, 6558-6562, 1982  
A:Title: Basic proline-rich proteins from human parotid saliva: complete covalent structure  
A:Reference number: A90464; MUID:80101329; PMID:6924859  
A:Contents: IB-9  
A:Accession: A90464  
A:Molecule type: protein  
A:Residues: 92-127,'R',129-148,'R',150-152 <KA3>  
R:Isemura, S.; Saitoh, E.; Sanada, K.  
J. Biochem. 91, 2067-2075, 1982

A>Title: Fractionation and characterization of basic proline-rich peptides of human parv  
A:Reference number: A91966; MUID:83007119; PMID:7118863  
A:Contents: P-E  
A:Accession: A91966  
A:Molecule type: protein  
A:Residues: 92-127, 'R', 129-148, 'R', 150-152 <ISS>  
C:Comment: this peptide contains 21-residue repeats, two of which have internal 7-residu  
C:Genetics:  
A:Gene: GDB: PRB1  
A:Cross-references: GDB:119511; OMIM:180989  
A:Map position: 12p13.2-12p13.2  
A:Note: each of the tandem repeats contains a candidate splice acceptor site, and severa  
C:Superfamily: proline-rich protein  
C:Keywords: alternative splicing; duplication; parotid gland; phosphoprotein; pyroglutam  
F:1-16/Domains: signal sequence #status predicted <SIG>  
F:17-51, 92-152/Product: basic proline-rich peptide IB-1 #status experimental <IB1>  
F:92-152/Product: basic proline-rich peptide P-E #status experimental <PPE>  
F:275-335/Product: basic proline-rich peptide IB-6 #status experimental <PIB6>  
F:337-392/Product: basic proline-rich peptide P-F #status experimental <PFF>  
F:337-392/Product: basic proline-rich peptide P-H #status experimental <PPH>  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen  
F:24/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 11.6%; Score 137.5; DB 1; Length 392;  
Best Local Similarity 26.6%; Pred. No. 0.042; Indels 71; Gaps 11;  
Matches 59; Conservative 17; Mismatches 75;

QY 11 RSSSTSGGGLTLGLEPSEEARSPPLAP---SEG---AGSDVFDGDLGMAKGLQSL 64  
DB 86 KSSRPSPPGK-PGQPPPGGNGPPGPPPPGKPGQPPGKPKPGQPPGPPGKPGQ--- 140  
QY 65 PHDPSPLQRYSDPTVPLPSETDGYVAPLTCSPOPEYVNDVPPPP----- 113  
DB 141 ---PPPGGKSGSPSPSP-PGKPGQ-----PPPGGNGPPGPPPPGKPGQPPGQ 187  
QY 114 SPREGFLPARPAGATLERAKTLT---DGKNGVAVKDVAFAGAVENPEVLTPOGGAAPQ 169  
DB 188 NKQGGPPGPPGKPGQPPGKSGSPSPSPGK-----PGQPPGNGQPP 231  
QY 170 PHPPAFSPAFDNLVYWDOPPERGA-----PSTFKGTP 204  
DB 232 GPPPPGKPP-----GGPPGNGRPPGPPPPGKPPG 264

RESULT 10  
H70786  
probable dihydroliipoamide acetyltransferase component [similarity] - Mycobacterium tuberc  
C:Species: Mycobacterium tuberculosis  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 01-Sep-2000  
C:Accession: H70786  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S  
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Genies, S.; Hamlin, N.; Holroyd, S.  
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: H70786  
A:Molecule type: DNA  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Residues: 1-553 <CON>  
A:Cross-references: GB:Z70283; GB:AL123456; MTD:93261561; PIDN:CAA94256.1; PID:91237068  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: sucB  
C:Superfamily: Mycobacterium probable dihydroliipoamide succinyltransferase; Lipoyl/bioti  
F:4-77/Domains: lipoyl/biotin-binding homology <LBBI>  
F:43/Binding site: liipoamide (Lys) (covalent) #status predicted  
F:162/Binding site: liipoamide (Lys) (covalent) #status predicted

Query Match 11.5%; Score 135.5; DB 1; Length 553;  
Best Local Similarity 25.9%; Pred. No. 0.083; Indels 61; Gaps 10;  
Matches 56; Conservative 24; Mismatches 75;

QY 27 EPSEEARSPPL--APSEGAGSDVFDGDLGMAKGLQSL-----L 61  
DB 101 KPAPEPPPPVQPTGACAPAGDAPKLVLPBELGESVTEGVIRMLKIGDSVQVDEPLEVEVS 160  
QY 62 QSLPTDPSPLQ-----KSEDPVPLPSETD-GYVAPLTCSPOPEYVNDVPPPP 112  
DB 161 DKYDTEIPSPVAVLVLSISADEDATVPGGELARIGVAAIDIGAAPAKPKPKP--VPEPA 218  
QY 113 PSPREPPL--PAARPAGA-----TLERAKTLSPGKNGVVK--DVFA 149  
DB 219 PTKAEPAPSPPAQAPAGAAEGAPVTPVLRKLASENNIDLAGVGTGTGGRIRKQDVLA 278  
QY 150 FGGAIVENPEVLP--OGGAAPQPHPPAPSPAFDNL 183  
DB 279 AAEQKKRAKAPAPAAQAAAPAPKAPAPAPALAH 314

RESULT 11  
S06733  
hydroxyproline-rich glycoprotein precursor - common tobacco  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 21-Jul-2000  
C:Accession: S06733  
R:Keller, B.; Lamb, C.J.  
Genes Dev. 3, 1639-1646, 1989  
A>Title: Specific expression of a novel cell wall hydroxyproline-rich glycoprotein ge  
A:Reference number: S06733; MUID:90128263; PMID:2612909  
A:Accession: S06733  
A:Molecule type: DNA  
A:Residues: 1-620 <KEU>  
A:Cross-references: EMBL:X13885; MTD:919866; PIDN:CAA32090.1; PID:919867  
C:Superfamily: hydroxyproline-rich glycoprotein  
C:Keywords: glycoprotein

Query Match 11.4%; Score 135; DB 2; Length 620;  
Best Local Similarity 26.9%; Pred. No. 0.1; Indels 82; Gaps 14;  
Matches 66; Conservative 18; Mismatches 79;

QY 18 GCGDLTLGLEPSEEARSPPL--APSEGAGSDVFDGDLGMAKGLQSLPTG-DPSPLQ 73  
DB 130 GPHNL-----PSHGQRPSPSHGHAPPSG-----GTPPRG-QNPPSRARRSPS 173  
QY 74 R-----YSDPTVPLPSETDGYVAPLTCSPOPEYVNDV-----RQPPSPRSG 118  
DB 174 RHGHPRPPTVAPQPPPTVSPSPQVOPRPPTVSPPTVHQPSPSPSKGHPQPPPTNRNA 233  
QY 119 PLPAARPAGATLE-----RAKTLSPGKNGVAVKDVAFAGAVENPEVLTPO 163  
DB 234 P-PTNHAPPTHQPSPPLRLPSPFRQPOPPTVSPPPAPAYQ-----SPQSPPTVSP 286  
QY 164 GGAAPQ-----HPPAFSPA-----FDNLVYWDOPPERGAPSTFKGTPFAENP 209  
DB 287 PTPSPPPSPPTVSPPPAPVSPSPPTPTPTPF-----SPPPAPVSPPTVSPPP-----P 335  
QY 210 EYVGL 214  
DB 336 TYVLP 340

RESULT 12  
S31719  
proline-rich protein - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Nov-1999  
C:Accession: S31719  
R:Nishimatsu, S.I.; Sato, K.Y.Z.; Oda, K.Y.Z.; Noto, K.Y.Z.; Ueno, K.Y.Z.  
submitted to the EMBL Data Library, August 1992  
A:Reference number: S31719  
A:Accession: S31719  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-379 <NIS>

A:Cross-references: EMBL:X68249; NID:g64955; PID:g64956  
C:Superfamily: proline-rich protein

Query Match 11.4%; Score 134.5; DB 2; Length 379;  
Best Local Similarity 28.08; Pred. No. 0.064;  
Matches 59; Conservative 16; Mismatches 97; Indels 39; Gaps 9;

QY 2 AGCVHHRHSSSTRSGGDLTLGLEPSEEARSPFLAPSEGG--AGSDVFDGLMGAA 58  
DB 134 SGGRRPHDDSDGNNRSPREYGRNHRSLPLSPSPSTSGMKHSSSAPPSPRRQA 193  
QY 59 KGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNDVAPRPPSP-- 115  
DB 194 -GAPPAASQNAKKRYNR--EKPLPTPGHRAAPVAPVPPSPINSRSPSAHSQPPPPPY 250  
QY 116 -----REGPLAARAGATLERAKTLSPKNGVAKDVFAGAVENPE--YITPOGA 166  
DB 251 RQPTSLNGRPPSPINEAPELPQRHNSLHRKTAGPVR---GLAPPPQSVHLSP--G 304  
QY 167 APQHPAPAFSPADNLYWMDPPERGAPP 197  
DB 305 RP---PPPA-----KQPPGKGAAP 320

RESULT 13  
T29293  
hypothetical protein C50F7.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T29293  
R:Johnson, D.; Stellyes, L.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C50F7.  
A:Reference number: 220601

A:Accession: T29293  
A:Status: preliminary; translated from GB/EMBL/DDBT  
A:Molecule type: DNA  
A:Residues: 1-309 <JOH>  
A:Cross-references: EMBL:U41557; PID:AAA83301.1; CESP:C50F7.5  
C:Genetics:  
A:Gene: CESP:C50F7.5  
C:Superfamily: unassigned collagens

Query Match 11.3%; Score 134; DB 2; Length 309;  
Best Local Similarity 24.2%; Pred. No. 0.055;  
Matches 54; Conservative 25; Mismatches 74; Indels 70; Gaps 10;

QY 7 HHRHRSSTNSGGDLTLGLEPSEEARSPFLAPSEGGSDVFDGLMGAAKGLQSLPT 66  
DB 31 HHHHKTAPRTSRGIATTTFAPTSSDL---PIAGSSSA----- 65  
QY 67 HDPSPLQRYSEDP--TYPLPSE--TDGVNAPLTCSPQPEYVNDVAPRPPSPREGP 119  
DB 66 ----PVTAASADPLPTSVVPQSPNEPSPGVAP--SDSPSPGPPS--PGVNPSEDP 116  
QY 120 LPAARPGATLERAKTLSPKNGVAKDVFAGAVENPEYITPOGAAPQ-----HPP 174  
DB 117 OPSPGPSPGPVDPSEDPQ-----SVEPSEMDQSPGPPSPGTVDPSEDPQ 162  
QY 175 AFSAPFNLYWMDPPERGAPPSTFGTPAENPEYLGLDVP 217  
DB 163 SVPESED-----PQSPGPPSPGPVDP--SEDPQSPGSSSP 195

RESULT 14  
T05264  
probable serine/threonine-specific protein kinase (EC 2.7.1.1) T4L20.20 - Arabidopsis th  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 17-Mar-2000  
C:Accession: T05264  
R:Bevan, M.; Terry, N.; Adlles, W.; Buysnaert, C.; Dasseville, R.; De Clerck, R.; De  
ewen, H.W.; Mayer, K.F.X.; Schueller, C.  
submitted to the Protein Sequence Database, September 1998

A:Reference number: Z15406  
A:Accession: T05264  
A:Molecule type: DNA  
A:Residues: 1-674 <BEV>  
A:Cross-references: EMBL:AL023094  
A:Experimental source: cultivar Columbia; BAC clone T4L20  
C:Genetics:  
A:Map position: 4  
A:Introns: 180/1; 221/1; 381/1; 421/1; 444/3; 470/2; 518/3; 583/3  
A:Note: T4L20.20  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
C:Keywords: phosphotransferase; protein kinase

Query Match 11.3%; Score 134; DB 2; Length 674;  
Best Local Similarity 23.9%; Pred. No. 0.13;  
Matches 59; Conservative 22; Mismatches 76; Indels 90; Gaps 12;

QY 17 SGGDLTLGLEPSEEARSPFLAPSEGGSDVFDGLMGAAKGLQSLPTHD----- 68  
DB 15 SNGTPSNGTSPSNSSPPRP--PSSPPSSI-----SAPPDIGASFSPP 58  
QY 69 PSPQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNDVAPRPPSPREG-----PLPA 122  
DB 59 PAPPTQETSPPTSPSSS-----PPVYANSP-----QTPENSPAP--EGSTPYTPAPP 107  
QY 123 ARPAGATLERAKTLSPKNGVAKDVFAGAVENPEYITPO-----GGAPOHPHP 174  
DB 108 QTPSNQSPERTPTSPGAN---DDRNTNGNNRNDSTSPSSGRTSDGSPSPPR 164  
QY 175 AF-----SPAFNLYWMDPPERGAPPST- 199  
DB 165 SISPQNSGDSSSGILLLLAVLCICICNKKKKKKSPQVHHMYNNN--PYGAPSGNG 223  
QY 200 --FKGTP 204  
DB 224 GYKGTGTP 230

RESULT 15  
A53183  
epidermal growth factor receptor precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999  
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643  
R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A  
Genes Dev. 8, 399-413, 1994  
A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF recepto  
A:Reference number: A53183; MUID:94170986; PMID:8125255  
A:Accession: A53183  
A:Molecule type: mRNA  
A:Residues: 1-1210 <LUE>  
A:Cross-references: GB:U03425  
R:Aviv, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.  
Oncogene 6, 673-676, 1991  
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding  
A:Reference number: A43818; MUID:91232866; PMID:2030916  
A:Accession: A43818  
A:Molecule type: mRNA  
A:Residues: 1-714 <AVI>  
R:Elisinger, D.P.; Serrero, G.  
submitted to the EMBL Data Library, June 1992  
A:Reference number: S24942  
A:Accession: S24942  
A:Molecule type: mRNA  
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>  
A:Cross-references: EMBL:Z12608  
R:Heisermann, G.J.; Gill, G.N.  
J. Biol. Chem. 263, 13152-13158, 1988  
A:Title: Epidermal growth factor receptor: threonine and serine residues phosphorylate  
A:Reference number: A28941; MUID:88350814; PMID:3138223  
A:Accession: A28941  
A:Molecule type: protein

A:Residues: 689-694,'X',696-704,'L',706-707;989-992,'XX',995-996,'X',998-1000;1002-1009,  
 R:Hiibs, M.L.; Dunn, A.R.; Alexander, W.S.  
 submitted to the EMBL Data Library, April 1994  
 A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor  
 A:Reference number: S45325  
 A:Accession: S45325  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-971,'K',973-1210 <VER>  
 A:Cross-references: EMBL:X78987; NID:9488830; PIDN:CAA55587.1; PID:9488831  
 R:Parls, B.C.; Das, S.K.; Andrews, G.K.; Day, S.K.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993  
 A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse b  
 A:Reference number: 149643; MUID:93126380; PMID:7678348  
 A:Accession: 149643  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 12-20,22-132 <RES>  
 A:Cross-references: GB:L06864; NID:9193001; PIDN:AAA53029.1; PID:9567201  
 C:Genetics:  
 A:Gene: EGFR  
 C:Superfamily: epidermal growth factor receptor; protein kinase homology  
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphop  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:648-670/Domain: transmembrane #status predicted <TM>  
 F:712-977/Domain: protein kinase homology <KIN>  
 F:720-728/Region: protein kinase ATP-binding motif  
 F:680-695/Binding site: phosphate (Thr) (covalent) #status experimental  
 F:697-1070/Binding site: phosphate (Ser) (covalent) #status experimental  
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental  
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental  
 F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 11.3%; Score 134; DB 2; Length 1210;  
 Best local Similarity 26.4%; Pred. No. 0.24;  
 Matches 58; Conservative 23; Mismatches 63; Indels 76; Gaps 11;

QY	47	DVF	DGLG	MAKGL	SLP	THD	PSP	-----	LQ	YS	76
DB	1011	DV	DAD	EY	L	P	O	G	F	N	S
QY	77	ED	P	T	V	L	P	S	E	T	-----
DB	1071	SD	P	T	G	A	V	T	E	N	I
QY	135	T	L	S	P	G	K	N	G	V	K
DB	1118	-	A	P	G	R	D	L	H	Q	N
QY	189	---	D	P	-----	P	E	R	G	A	
DB	1165	M	S	L	D	N	P	D	Y	O	

Search completed: April 28, 2003, 13:42:33  
 Job time : 11.9833 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:08 ; Search time 17.4181 Seconds

(without alignments)  
1660.081 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182  
Sequence: 1 GAGGMVHRRSRSTRSGG.....STFKGPTAENPEYLGDVP 217

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182	100.0	217	22	AAE13121
2	1182	100.0	397	22	AAE13122
3	1182	100.0	555	22	AAE13108
4	1182	100.0	564	22	AAE13110
5	1182	100.0	690	22	AAE13109
6	1182	100.0	697	22	AAE13111
7	1182	100.0	1179	22	AAE13123
8	1182	100.0	1223	23	AAU98923
9	1182	100.0	1255	21	AAV92620
10	1182	100.0	1255	22	AAE12130

11	1182	100.0	1255	22	AAE60167
12	1182	100.0	1255	23	AAU74545
13	1177	99.6	266	21	AAE21201
14	1177	99.6	266	23	AAE51146
15	1177	99.6	583	23	AAE20483
16	1177	99.6	587	23	AAE20481
17	1177	99.6	587	23	AAE20484
18	1177	99.6	600	23	AAE20482
19	1177	99.6	919	21	AAE21203
20	1177	99.6	919	23	AAE51148
21	1177	99.6	1255	17	AAW01111
22	1177	99.6	1255	20	AAW92406
23	1177	99.6	1255	21	AAE21158
24	1177	99.6	1255	21	AAE20479
25	1177	99.6	1255	22	AAE20478
26	1177	99.6	1255	22	AAE20479
27	1177	99.6	1255	23	AAE20479
28	1177	99.6	1255	23	AAE20479
29	1177	99.6	1255	23	AAE20479
30	1177	99.6	1255	23	AAE20479
31	1177	99.6	1433	14	AAU77114
32	1086	91.9	287	19	AAE39568
33	961	81.3	920	23	AAE51133
34	961	81.3	926	23	AAE51152
35	961	81.3	1256	21	AAE21206
36	961	81.3	1256	22	AAE21206
37	961	81.3	1256	23	AAE21206
38	954	80.7	1256	21	AAE21206
39	954	80.7	1256	23	AAE21206
40	864	73.1	1200	21	AAE21208
41	711	60.2	293	21	AAE58875
42	648.5	54.9	135	21	AAE3787
43	164	13.9	1210	21	AAE19259
44	164	13.9	1210	21	AAE50616
45	164	13.9	1210	22	AAE68420

#### ALIGNMENTS

RESULT 1  
AAE13121  
AAE13121 standard; Protein: 217 AA.  
AC AAE13121;  
XX 28-JAN-2002 (first entry)  
DE Mature human HER-2 membrane distal intracellular domain.  
XX  
XX  
XX  
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
KW human; HER-2 membrane distal intracellular domain.  
XX  
XX Homo sapiens.  
XX  
XX WO200174855-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US10515.  
XX  
XX 30-MAR-2000; 2000US-193504P.  
XX  
XX (DEND-) DENDREON CORP.  
XX  
XX Laus R, Vidovic D, Graddis T;  
XX WPT: 2001-662965/76.  
XX N-PSDB: AAD21572.  
XX  
XX An immunostimulatory fusion protein comprising the intracellular domain

HER2 transgene pla  
Human HER2 (ErbB2)  
Human HER-2/neu pr  
Human HER-2/neu on  
Human protein for  
Human protein for  
Human protein for  
Human protein for  
Human protein for  
Human HER-2/neu fu  
Her-2/neu extracel  
HER-2/neu protein.  
Human HER-2/neu on  
Human HER-2/neu pr  
Amino acid sequenc  
Human HER-2/neu pr  
HER2/neu amino aci  
Human Her-2 protei  
Human Her-2/neu pr  
Human Her-2/neu on  
Human Her-2/neu po  
Sequence of c-erbB  
HER-2/neu polypept  
Human Her-2/neu ex  
Mouse Her-2/neu ex  
Mouse Her-2/neu pr  
Amino acid sequenc  
Mouse Her-2/neu on  
Rat Her-2/neu prot  
Rat Her-2/neu onco  
Human HER-2/neu pr  
Breast and ovarian  
Human cancer assoc  
Amino acid sequenc  
Human EGF receptor  
Amino acid sequenc

PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 2; Page 31; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is  
 CC mature human HER-2 membrane distal intracellular domain.  
 CC This sequence is used in the HER500 and HER300 GM-CSF fusion constructs  
 CC of the invention.  
 CC  
 XX

SQ Sequence 217 AA;

Query Match 100.0%; Score 1182; DB 22; Length 217;

Best Local Similarity 100.0%; Pred. No. 1.8e-91;

Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMAK 60  
 DB 1 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMAK 60  
 QY 61 LQSLPTHPDLRYSDDPVLPSEETDGYVAFLTCSPOEYVNOPVRRQPPSPRGPL 120  
 DB 61 LQSLPTHPDLRYSDDPVLPSEETDGYVAFLTCSPOEYVNOPVRRQPPSPRGPL 120  
 QY 121 PAARPGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGAAPQHPPPAFSPAF 180  
 DB 121 PAARPGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGAAPQHPPPAFSPAF 180  
 QY 181 DNLYYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217  
 DB 181 DNLYYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217

RESULT 2

AAE13122 standard; Protein; 397 AA.

AC AAE13122;

DT 28-JAN-2002 (first entry)

DE NY-ESO-1C-HER-2 membrane distal intracellular domain fusion protein.

KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW human; HER-2 membrane distal intracellular domain; NY-ESO-1C;  
 KW autoimmunogenic cancer/testis antigen.

OS Homo sapiens.

PN WO200174855-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US10515.

PR 30-MAR-2000; 2000US-193504P.

PA (DEND-) DENDREON CORP.

PI Laus R, Vidovic D, Graddis T;

DR WPI; 2001-662965/76.

PI N-PSDB; AAD21573.

PT An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Disclosure; Page 54-55; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is a  
 CC fusion protein which comprises human autoimmunogenic  
 CC cancer/testis antigen, NY-ESO-1C and mature human HER-2 membrane distal  
 CC intracellular domain.  
 CC  
 XX

SQ Sequence 397 AA;

Query Match 100.0%; Score 1182; DB 22; Length 397;

Best Local Similarity 100.0%; Pred. No. 3.6e-91;

Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMAK 60  
 DB 181 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMAK 240  
 QY 61 LQSLPTHPDLRYSDDPVLPSEETDGYVAFLTCSPOEYVNOPVRRQPPSPRGPL 120  
 DB 241 LQSLPTHPDLRYSDDPVLPSEETDGYVAFLTCSPOEYVNOPVRRQPPSPRGPL 300  
 QY 121 PAARPGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGAAPQHPPPAFSPAF 180  
 DB 301 PAARPGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGAAPQHPPPAFSPAF 360  
 QY 181 DNLYYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217  
 DB 361 DNLYYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 397

RESULT 3

AAE13108 standard; Protein; 555 AA.

AC AAE13108;

DT 28-JAN-2002 (first entry)

DE Human HER500 fusion protein construct.

KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal intracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human;  
 KW HER-2 protein; HER500 fusion protein.

OS Chimeric - Homo sapiens.

PN WO200174855-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US10515.

PR 30-MAR-2000; 2000US-193504P.

PA (DEND-) DENDREON CORP.

PI Laus R, Vidovic D, Graddis T;

DR WPI: 2001-662965/76.  
 DR N-PSDB: AAD21564.  
 XX  
 PT An immunostimulatory fusion protein comprising the intracellular domain  
 of HER-2 and an antigen elicits an immune response to the antigen and  
 is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 26; 59pp; English.  
 XX  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC fusion protein construct which comprises human PAP signal  
 CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal  
 CC sequence, mature HER-2 membrane distal extracellular and intracellular  
 CC domains and a C-terminal tag.  
 CC  
 XX  
 SQ Sequence 555 AA;  
 Query Match 100.0%; Score 1182; DB 22; Length 555;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-91;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAGGMVHHRHSSSTRSGGDLTGLPSEEARSPPLAPSEAGSDVFDGLMGAAKG 60  
 DB 330 GAGGMVHHRHSSSTRSGGDLTGLPSEEARSPPLAPSEAGSDVFDGLMGAAKG 389  
 QY 61 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRSGPL 120  
 DB 390 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRSGPL 449  
 QY 121 PAARPAGATLERAKTSLPGKNGVYKDYAFAGAVENPEYLTPOGGAAPQHPAPAFSPAF 180  
 DB 450 PAARPAGATLERAKTSLPGKNGVYKDYAFAGAVENPEYLTPOGGAAPQHPAPAFSPAF 509  
 QY 181 DNLTYWDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
 DB 510 DNLTYWDPPERGAPSTFKGTPTAENPEYLGIDVP 546  
 RESULT 4  
 ID AAE13110 standard; Protein: 564 AA.  
 AC AAE13110;  
 XX  
 DT 28-JAN-2002 (first entry)  
 XX  
 DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.  
 XX  
 KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human; OVA;  
 KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Unidentified.  
 XX  
 PN WO200174855-A2.  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001KO-US10515.  
 XX  
 PR 30-MAR-2000; 2000US-193504P.  
 XX

PA (DEND-) DENDREON CORP.  
 XX  
 PI Laus R, Vidovic D, Graddis T;  
 XX  
 DR WPI: 2001-662965/76.  
 DR N-PSDB: AAD21566.  
 XX  
 PT An immunostimulatory fusion protein comprising the intracellular domain  
 of HER-2 and an antigen elicits an immune response to the antigen and  
 is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 26; 59pp; English.  
 XX  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC fusion protein construct which comprises human PAP signal  
 CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal  
 CC sequence, mature HER-2 membrane distal extracellular domain, an  
 CC Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,  
 CC HER-2 membrane distal intracellular domain and a C-terminal tag.  
 CC  
 XX  
 SQ Sequence 564 AA;  
 Query Match 100.0%; Score 1182; DB 22; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-91;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAGGMVHHRHSSSTRSGGDLTGLPSEEARSPPLAPSEAGSDVFDGLMGAAKG 60.  
 DB 339 GAGGMVHHRHSSSTRSGGDLTGLPSEEARSPPLAPSEAGSDVFDGLMGAAKG 398  
 QY 61 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRSGPL 120  
 DB 399 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRSGPL 458  
 QY 121 PAARPAGATLERAKTSLPGKNGVYKDYAFAGAVENPEYLTPOGGAAPQHPAPAFSPAF 180  
 DB 459 PAARPAGATLERAKTSLPGKNGVYKDYAFAGAVENPEYLTPOGGAAPQHPAPAFSPAF 518  
 QY 181 DNLTYWDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
 DB 519 DNLTYWDPPERGAPSTFKGTPTAENPEYLGIDVP 555  
 RESULT 5  
 ID AAE13109 standard; Protein: 690 AA.  
 AC AAE13109;  
 XX  
 DT 28-JAN-2002 (first entry)  
 XX  
 DE Human HER500-hgm-CSF fusion protein construct.  
 XX  
 KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;  
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;  
 KW HER500-hgm-CSF fusion protein.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Synthetic.  
 XX  
 PN WO200174855-A2.  
 XX

PD 11-OCT-2001.  
 XX 30-MAR-2001; 2001MO-US10515.  
 PF 30-MAR-2000; 2000US-193504P.  
 PR (DEND-) DENDREON CORP.  
 XX  
 XX Laus R, Vidovic D, Graddis T;  
 DR WPI: 2001-662965/76.  
 N-PSDB; AAD21565.  
 XX  
 XX An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 26; 59pp: English.  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC hGM-CSF fusion protein construct which comprises human PAP  
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2  
 CC signal sequence, mature HER-2 membrane distal extracellular and  
 CC intracellular domains, an Ala Ala linker, a mature human granulocyte-  
 CC macrophage colony stimulating factor (GM-CSF) sequence and a  
 CC C-terminal tag.  
 CC  
 XX  
 SQ Sequence 690 AA:  
 Query Match 100.0%; Score 1182; DB 22; Length 690;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-91;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFGDLMGAAKG 60  
 DB 330 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFGDLMGAAKG 389  
 QY 61 LQSLPTHDPSPFLQRYSDPVPPLPSETDGYVAPLTGSPQPEYVNPQPVRRQPSPRGCL 120  
 DB 390 LQSLPTHDPSPFLQRYSDPVPPLPSETDGYVAPLTGSPQPEYVNPQPVRRQPSPRGCL 449  
 QY 121 PAARPAGATLERAKTISPGKNGVVKDVFARFGAVENPEYLTPOGGAAPQHPPPAASPAF 180  
 DB 450 PAARPAGATLERAKTISPGKNGVVKDVFARFGAVENPEYLTPOGGAAPQHPPPAASPAF 509  
 QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 217  
 DB 510 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 546  
 RESULT 6  
 AAE13111  
 ID AAE13111 standard; Protein: 697 AA.  
 XX  
 AC AAE13111;  
 DT 28-JAN-2002 (first entry)  
 XX  
 DE Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.  
 XX  
 KW Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;  
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;  
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;

KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Rattus norvegicus.  
 OS Chimeric - Unidentified.  
 XX  
 XX WO200174855-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001MO-US10515.  
 DR 30-MAR-2000; 2000US-193504P.  
 PR (DEND-) DENDREON CORP.  
 XX  
 XX Laus R, Vidovic D, Graddis T;  
 PI WPI: 2001-662965/76.  
 DR N-PSDB; AAD21567.  
 XX  
 XX An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX  
 PS Claim 7; Page 27; 59pp: English.  
 CC The invention relates to immunostimulatory fusion proteins (IFP) and  
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 CC polypeptide antigen component and an immunostimulatory component derived  
 CC from the intracellular domain of HER-2 protein which is effective to  
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 CC immune response to the antigen. IFP or superactivated dendritic cells  
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer.  
 CC associated with a particularly antigen. The present sequence is HER500  
 CC rGM-CSF fusion protein construct which comprises human PAP  
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2  
 CC signal sequence, mature HER-2 membrane distal extracellular domain,  
 CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,  
 CC HER-2 membrane distal intracellular domain, an Ala Ala linker, a mature  
 CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence  
 CC and a C-terminal tag.  
 CC  
 XX  
 SQ Sequence 697 AA:  
 Query Match 100.0%; Score 1182; DB 22; Length 697;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-91;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFGDLMGAAKG 60  
 DB 339 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFGDLMGAAKG 398  
 QY 61 LQSLPTHDPSPFLQRYSDPVPPLPSETDGYVAPLTGSPQPEYVNPQPVRRQPSPRGCL 120  
 DB 399 LQSLPTHDPSPFLQRYSDPVPPLPSETDGYVAPLTGSPQPEYVNPQPVRRQPSPRGCL 448  
 QY 121 PAARPAGATLERAKTISPGKNGVVKDVFARFGAVENPEYLTPOGGAAPQHPPPAASPAF 180  
 DB 459 PAARPAGATLERAKTISPGKNGVVKDVFARFGAVENPEYLTPOGGAAPQHPPPAASPAF 518  
 QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 217  
 DB 519 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 555  
 RESULT 7  
 AAE13123  
 ID AAE13123 standard; Protein: 1179 AA.  
 XX  
 AC AAE13123;  
 DT 28-JAN-2002 (first entry)



XX DE Human SART3-IC-HER-2 membrane distal intracellular domain fusion protein.  
 XX DE Immunostimulatory fusion protein; IFP; antigen component; therapy;  
 KW immunostimulatory component; T-cell mediated immune response; DC;  
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;  
 KW human; HER-2 membrane distal intracellular domain; SART3-IC;  
 KW squamous cell carcinoma antigen; fusion protein.  
 XX OS Homo sapiens.  
 XX OS WO200174855-A2.  
 XX PN 11-OCT-2001.  
 XX PD 30-MAR-2001; 2001WO-US10515.  
 XX PE 30-MAR-2000; 2000US-193504P.  
 XX PR 30-MAR-2000; 2000US-193504P.  
 XX PA (DENND-) DENDREON CORP.  
 XX PI Laus R, Vidovic D, Graddis T;  
 XX DR WPI: 2001-662965/76.  
 XX DR N-PSDB: AAD21574.  
 XX XX  
 PT An immunostimulatory fusion protein comprising the intracellular domain  
 PT of HER-2 and an antigen elicits an immune response to the antigen and  
 PT is useful for the treatment of associated cancer associated -  
 XX XX  
 XX Disclosure: Page 55-58; 59pp; English.  
 XX PS The invention relates to immunostimulatory fusion proteins (IFP) and  
 XX CC nucleic acid molecules encoding such proteins. The IFPs comprise a  
 XX CC polypeptide antigen component and an immunostimulatory component derived  
 XX CC from the intracellular domain of HER-2 protein which is effective to  
 XX CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular  
 XX CC immune response to the antigen. IFP or superactivated dendritic cells  
 XX CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer  
 XX CC associated with a particularly antigen. The present sequence is a  
 XX CC fusion protein which comprises human squamous cell  
 XX CC carcinoma antigen, SART3-IC and mature human HER-2 membrane distal  
 XX CC intracellular domain.  
 XX CC  
 SO Sequence 1179 AA;  
 Query Match 100.0%; Score 1182; DB 22; Length 1179;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-90;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GAGGWHHRHSSSTRSGGDLTLGLEPSEEBAPRSLAPSBGASDVFDGDLGGAAG 60  
 Db 963 GAGGWHHRHSSSTRSGGDLTLGLEPSEEBAPRSLAPSBGASDVFDGDLGGAAG 1022  
 Oy 61 LQSLPTHDPSPLOQRYSEDPYPLPSETDGYVAPLTCSPOPEYVNPQDVRRPSPREGPL 120  
 Db 1023 LQSLPTHDPSPLOQRYSEDPYPLPSETDGYVAPLTCSPOPEYVNPQDVRRPSPREGPL 1082  
 Oy 121 PAARAGATLERAKTSLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
 Db 1083 PAARAGATLERAKTSLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 1142  
 Oy 181 DNLTYWDDPPERGAPSTFGKTPAENPEYLTGLDVP 217  
 Db 1143 DNLTYWDDPPERGAPSTFGKTPAENPEYLTGLDVP 1179

DT 24-SEP-2002 (first entry)  
 XX DE Human breast cancer antigen, Her2 variant.  
 XX DE Human; Her2; cyrostatic; antiviral; immunostimulant;  
 KW cell-mediated immune response; tumour; breast cancer;  
 KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;  
 KW lymphoma; leukaemia; hepatitisvirus; lentivirus; herpesvirus;  
 KW human immunodeficiency virus; HIV; flavivirus; pestivirus.  
 XX OS Homo sapiens.  
 XX OS WO200240059-A2.  
 XX PN 23-MAY-2002.  
 XX PD 01-NOV-2001; 2001WO-US45626.  
 XX PE 01-NOV-2000; 2000US-0704232.  
 XX PR 01-NOV-2000; 2000US-0704232.  
 XX PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.  
 XX PA (MINC/) MINCHEFF M S.  
 XX PA (LOUK/) LOUKINOV D I.  
 XX PA (ZOUB/) ZOUBAK S.  
 XX PI Mincheff MS, Loukinov DI, Zoubak S;  
 XX DR WPI: 2002-527524/56.  
 XX DR N-PSDB: ABK86207.  
 XX XX  
 PT Inducing a cell-mediated immune response against a target antigen, by  
 PT reducing undesired cells and stimulating presentation of an antigen by  
 PT a cell, comprises administering a polynucleotide encoding a variant of  
 PT an antigen -  
 XX XX  
 XX Claim 21; Page 134-138; 146pp; English.  
 XX PS The invention relates to a method of inducing a cell-mediated immune  
 XX CC response against a cell comprising a target antigen (I) in a subject,  
 XX CC treating a subject having undesired cells, for example tumour cells  
 XX CC or vitally infected cells (C), reducing the number of (C) in a subject,  
 XX CC and stimulating presentation of (I) by a cell. This is done by  
 XX CC administering a polynucleotide (II) encoding a variant of (I), so that  
 XX CC (II) expressed in a cell and cell-mediated immune response is induced.  
 XX CC The method can be used to treat prostate cancer, breast cancer,  
 XX CC colorectal cancer and pancreatic cancer, as well as lymphomas and  
 XX CC leukaemias. The method is also useful in treating chronic viral  
 XX CC infections such as those caused by hepatitisviruses, lentiviruses and the  
 XX CC (including human immunodeficiency virus (HIV)), herpesviruses and the  
 XX CC human breast cancer antigen, Her2 variant, used as a target  
 XX CC antigen in the method of the invention.  
 XX CC  
 SO Sequence 1223 AA;  
 Query Match 100.0%; Score 1182; DB 23; Length 1223;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-90;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GAGGWHHRHSSSTRSGGDLTLGLEPSEEBAPRSLAPSBGASDVFDGDLGGAAG 60  
 Db 1006 GAGGWHHRHSSSTRSGGDLTLGLEPSEEBAPRSLAPSBGASDVFDGDLGGAAG 1065  
 Oy 61 LQSLPTHDPSPLOQRYSEDPYPLPSETDGYVAPLTCSPOPEYVNPQDVRRPSPREGPL 120  
 Db 1066 LQSLPTHDPSPLOQRYSEDPYPLPSETDGYVAPLTCSPOPEYVNPQDVRRPSPREGPL 1125  
 Oy 121 PAARAGATLERAKTSLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
 Db 1126 PAARAGATLERAKTSLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 1165  
 Oy 181 DNLTYWDDPPERGAPSTFGKTPAENPEYLTGLDVP 217

Db 1186 DNLTYMDPPPERGAPPSTFKGTPTAENPEYLGLDVP 1222

RESULT 9  
AAV92620  
ID AAV92620 standard; Protein; 1255 AA.  
XX AAV92620;  
XX 10-AUG-2000 (first entry)  
XX Human heregulin 2 (Her2).  
XX  
XX Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;  
KW self-protein; cancer; breast cancer; prostate cancer;  
KW cell-associated peptide antigen; foreign epitope.  
XX  
XX Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT 1..173 /label= N-terminal  
FT /note= "mature polypeptide"  
FT 5..25  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 59..73  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 103..117  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 149..163  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 174..323  
FT /label= Cysteine-rich\_domain  
FT 210..224  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 250..264  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 324..483  
FT /label= Ligand\_binding\_domain  
FT 325..339  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 369..383  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 465..479  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 484..623  
FT /label= Cysteine-rich\_domain  
FT 579..593  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 624..654  
FT /label= Transmembrane\_domain  
FT 632..652  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 653..667  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 655..1010  
FT /label= Tyrosine\_kinase\_domain  
FT 661..675  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 695..709

FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 710..730  
FT /label= insertion\_region  
FT /note= "suitable for foreign epitope insertion"  
FT 1011..1235  
FT /label= C-terminal\_domain  
XX  
XX WO200020027-A2.  
XX 13-APR-2000.  
XX  
XX 05-OCT-1999; 99NO-DK0525.  
XX  
XX 05-OCT-1998; 98DK-0001261.  
XX 20-OCT-1998; 98US-0105011.  
XX  
XX (MEBT-) M & E BIOTECH AS.  
XX  
XX Steinae L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;  
PI Gautam A, Birk P, Karlsson G;  
XX  
XX WPI: 2000-349917/30.  
XX N-PSDB; AAA09455.  
XX  
XX Inducing immune responses to weakly immunogenic, tumor associated  
PT peptide antigens for the treatment of breast and prostate cancer  
XX  
XX Claim 62; Page 193-198; 220pp; English.  
XX  
XX This is the human heregulin 2 (Her2) sequence. Immunogenic analogues of  
CC Her2 can be used in the claimed method as an autovaccine to induce a CTL  
CC response. Subdominant CTL epitopes, antibody binding regions and  
CC cysteine residues involved in disulfide bonds are preserved in the  
CC immunogenized forms. Regions suitable for the insertion of foreign T  
CC helper epitopes were identified (see features table). The method  
CC is used for inducing immune responses against weakly immunogenic  
CC cell-associated peptide antigens (PA) such as those associated with  
CC cancers (self-proteins), e.g. human prostate specific membrane antigen  
CC (PSM), heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b).  
CC The method comprises effecting simultaneous presentation by antigen  
CC producing cells (APCs) of the animals immune system of: (1) at least 1  
CC CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1  
CC B-cell group derived from the cell-associated PA; and (2) at least 1  
CC first T helper cell group which is foreign to the animal. Analogues of  
CC human PSM, human Her2 and human/murine FGF8b comprising a substantial  
CC part of all known and predicted CTL and B-cell epitopes of the respective  
CC PA and including at least one foreign T helper epitope are also claimed.  
CC The method is used to treat prostate, prostate/breast or breast cancer  
CC when the PA is human PSM, FGF8b and Her2, respectively.  
XX  
XX  
SQ Sequence 1255 AA:  
Query Match 100.0%; Score 1182; DB 21; Length 1255;  
Best Local Similarity 100.0%; Pred. NO. 1.3e-90;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAGGMVHRRSSSTRSGGDLTLGLEPSEEAAPRSLAPSEAGSDVPDGLGMGAAG 60  
DB 1038 GAGGMVHRRSSSTRSGGDLTLGLEPSEEAAPRSLAPSEAGSDVPDGLGMGAAG 1097  
QY 61 LQSLPHTDPSPLQRYSEDPTVPLPSETDGYVAPLTGSPQPEYVQNPVVRQPPSPRGPL 120  
DB 1098 LQSLPHTDPSPLQRYSEDPTVPLPSETDGYVAPLTGSPQPEYVQNPVVRQPPSPRGPL 1157  
QY 121 PAARPGATLEARKTSLSPGNGVVKDFAFGAVENTEYLTPOGGAAPQHPAPSPAF 180  
DB 1158 PAARPGATLEARKTSLSPGNGVVKDFAFGAVENTEYLTPOGGAAPQHPAPSPAF 1217  
QY 181 DNLTYMDPPPERGAPPSTFKGTPTAENPEYLGLDVP 217  
DB 1218 DNLTYMDPPPERGAPPSTFKGTPTAENPEYLGLDVP 1254

RESULT 10  
AAE12130 standard; Protein: 1255 AA.  
ID AAE12130 standard; Protein: 1255 AA.  
XX  
AC AAE12130;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human tyrosine kinase-type receptor, HER-2.  
XX  
XX Therapeutic compound; major histocompatibility complex; vaccine;  
KW antigenic peptide; MHC; immunoregulatory; immune response; HER-2;  
KW adoptive immunotherapy; anti-cancer; breast cancer antigen; APC;  
KW antigen presenting cell; human; tyrosine kinase-type receptor.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Region 774..782  
FT /note= "Antigenic epitope"  
XX  
FN WO200168677-A2.  
XX  
PD 20-SEP-2001.  
XX  
PE 16-MAR-2001; 2001WO-US40328.  
XX  
PR 16-MAR-2000; 2000US-0527487.  
XX  
PA (GEN2 ) GENZYME CORP.  
XX  
PI Nicotlette CA;  
XX  
DR WPI: 2001-616284/71.  
DR N-PSDB: AAD19731.  
XX  
XX  
PT Novel synthetic therapeutic compound for inducing immune response and  
PT for use in adoptive immunotherapy, has enhanced binding to major  
PT histocompatibility molecules and enhanced immunoregulatory properties  
XX  
XX  
PS Claim 4; Page 63-67; 69pp; English.  
XX  
XX The invention relates to synthetic therapeutic compounds (antigenic  
CC peptides) with enhanced binding to major histocompatibility complex  
CC (MHC) molecules and enhanced immunoregulatory properties relative  
CC to their natural counterparts. Compounds of the invention are useful  
CC for inducing an immune response in a subject and for use in adoptive  
CC immunotherapy. They are useful as components of anti-cancer vaccines  
CC and to expand immune effector cells that are specific for cancers  
CC characterised by expression of the breast cancer antigen, HER-2.  
CC Polynucleotides that encode peptides of the invention are useful as  
CC hybridisation probes and as primers for the detection of genes of gene  
CC transcripts that are expressed in antigen presenting cells (APCs), to  
CC confirm transduction of polynucleotides into host cells. The present  
CC sequence is human tyrosine kinase-type receptor, HER-2. Compounds  
CC of the invention are designed based on the HER-2 antigenic peptide  
CC (774-782).  
XX  
XX  
SQ Sequence 1255 AA;  
XX  
Query Match 100.0%; Score 1182; DB 22; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.3e-90;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 PAAPAGATLERAKTSPGKNGVYVDFAFGGAENPEYLPPOGGAAPQHPHPPAFSPAF 180  
DB 1158 PAAPAGATLERAKTSPGKNGVYVDFAFGGAENPEYLPPOGGAAPQHPHPPAFSPAF 1217  
QY 181 DNLVYWDODPPERGAPSTFGKTPTAENPEYLGIDVP 217  
DB 1218 DNLVYWDODPPERGAPSTFGKTPTAENPEYLGIDVP 1254  
RESULT 11  
AAB60167 standard; Protein: 1255 AA.  
ID AAB60167 standard; Protein: 1255 AA.  
XX  
AC AAB60167;  
XX  
DT 03-APR-2001 (first entry)  
XX  
DE HER2 transgene plasmid construct encoded protein.  
XX  
KW Human: HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;  
KW antibody.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
FN WO200100244-A2.  
XX  
PD 04-JAN-2001.  
XX  
PE 23-JUN-2000; 2000WO-US17229.  
XX  
PR 25-JUN-1999; 99US-0141316.  
PR 16-MAR-2000; 2000US-0189844.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Erickson S, Schwall R;  
XX  
DR WPI: 2001-061962/07.  
DR N-PSDB: AAF24297.  
XX  
XX  
PT Treating tumors, particularly breast cancers, which overexpress an ErbB  
PT receptor and does not respond to an anti-ErbB antibody, comprises  
PT conjugating the antibody to a maytansinoid -  
XX  
XX  
PS Example 3; Fig 4; 92pp; English.  
XX  
XX The present invention provides a method of treating cancer by  
CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In  
CC particular, the antibody is directed against ErbB2 (also known as HER2  
CC and p185neu). The method is particularly useful in the treatment of  
CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,  
CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.  
XX  
XX  
SQ Sequence 1255 AA;  
XX  
Query Match 100.0%; Score 1182; DB 22; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 1.3e-90;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

0Y      181 DNLVYWDOPPERGAPPSFTFKGTPTAENEPEYLGLDVP 217
        |||||||
DB      1218 DNLVYWDDOPPERRGAPPSTFKGTPTAENEPEYLGLDVP 1254

RESULT 12
AAU74545
ID      AAU74545 standard; Protein; 1255 AA.
AC      AAU74545;
DT      23-APR-2002 (first entry)
DE      Human HER2 (ErbB2) polypeptide.
KW      Human; HER2; ErbB; epidermal growth factor receptor; receptor;
KW      anti-ErbB antibody-maytansinoid conjugate; cancer; tumour; breast; ovary;
KW      stomach; endometrium; salivary gland; lung; kidney; colon; colorectum;
KW      thyroid; pancreas; prostate; bladder; ErbB2; neuronal disorder;
KW      glial disorder; astrocytal disorder; hypothalamic disorder;
KW      glandular disorder; macrophagal disorder; epithelial disorder;
KW      stromal disorder; blastocoealic disorder; inflammatory disorder;
KW      angiogenic disorder; immunological disorder.
XX      Homo sapiens.
OS      US2002001587-A1.
PN      03-JAN-2002.
PD      16-MAR-2001; 2001US-0811123.
PF      16-MAR-2000; 2000US-189844P.
PR      05-OCT-2000; 2000US-238327P.
XX      (ERIC/) ERICKSON S.
PA      (SCHW/) SCHWALL R.
PA      (SLIW/) SLIKOWSKI M.
XX      Erickson S, Schwall R, Slukowski M;
PI      WPI; 2002-163686/21.
DR      N-PSDB; ABK14058.
PT      Treating tumour characterised by overexpression of epidermal growth
XX      factor receptor, ErbB or cancer in mammal, comprises administering
XX      anti-ErbB antibody-maytansinoid conjugate to the mammal
XX      Example 3; Fig 7; 93pp; English.
PS      The invention relates to treating a tumour in a mammal, where the tumour
XX      is characterised by the overexpression of an epidermal growth factor
XX      receptor (ErbB) and does not respond or responds poorly, to treatment
XX      with an anti-ErbB antibody, comprising administering to the mammal an
XX      anti-ErbB antibody-maytansinoid conjugate. The method is useful for
XX      treating cancer or tumours of the breast, ovary, stomach, endometrium,
XX      salivary gland, lung, kidney, colon, colorectum, thyroid, pancreas,
XX      prostate and bladder, preferably breast cancer. The breast cancer is a
XX      metastatic breast cancer or an aggressive form of metastatic breast
XX      cancer which overexpresses ErbB2. The method is also useful for treating
XX      neuronal, glial, astrocytal, hypothalamic, glandular, macrophagal,
XX      epithelial, stromal, blastocoealic, inflammatory, angiogenic and
XX      immunological disorders. This sequence represents the human HER2 (ErbB2)
XX      polypeptide of the invention.
SO      Sequence 1255 AA;

Query Match          100.0%; Score 1182; DB 23; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1,3e-90;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0
1 GAGCGVVHHRHSSTRSGGDLTTLGLEPSEEARPSPLAPSGAGSDVDGDLGMKAAG 60
|||||||

```

```

Db 1038 GAGGWNHHNRSSSTRSGGDLTLGLPEEEEARSPFLAPSGAGSDVFDGGLGMAAG 1097
Qy 61 LOSLTPHDSPLQRISEDPTVLPSTBDYVAPLTCSFQPEVYNQPDVAPQPPSPREGUL 120
Db 1098 LOSLTPHDSPLQRISEDPVLPSTBDYVAPLTCSFQPEVYNQPDVAPQPPSPREGUL 1157
Qy 121 PARPAGATLEBAKTLSPKNGVWKVFAFGAVENPEYLTGOGGAAPQHPHPAPSPAF 180
Db 1156 PARPAGATLEBAKTLSPKNGVWKVFAFGAVENPEYLTGOGGAAPQHPHPAPSPAF 1217
Qy 181 DNLVYWDODPPERGAPPTSTFKGTPPAENPEYLGLDVP 217
Db 1218 DNLVYWDODPPERGAPPTSTFKGTPPAENPEYLGLDVP 1254

RESULT 13
AAB21201
ID AAB21201 standard; protein: 266 AA.
XX
AC AAB21201;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu protein phosphorylation domain.
XX
KM Human; HER-2/neu: oncogene; tyrosine kinase; cytostatic; vaccine.
KM breast cancer; prostate cancer; ovarian cancer; lung cancer;
KM colon cancer.
XX
OS Homo sapiens.
XX
PN WO200044899-A1.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
PA (CORI-) CORIXA CORP.
XX
PA (SMIK) SMITHKLINE BEECHAM.
XX
PI Cheever MA, Gheysen D;
XX
PI WPI: 2000-505976/45.
XX
DR HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Clam2; Flg 10; 128pp; English.
XX
CC The present sequence is the phosphorylation domain of the HER-2/neu
CC protein. HER-2/neu is a member of the tyrosine kinase family of
CC receptor-like glycoproteins and shows homology to the epidermal growth
CC factor receptor (EGFR). It probably plays a part in cell-growth and/or
CC differentiation. The HER-2/neu gene is an oncogene. An HER-2/neu fusion
CC protein comprising a HER-2/neu extracellular domain fused to a HER-2/neu
CC phosphorylation domain may be used to treat or prevent cancer by
CC eliciting or enhancing an immune response to the HER-2/neu protein. It
CC may be used to treat malignancies such as breast, ovarian, colon, lung
CC and prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 266 AA;

Query Match 99.6%; Score 1177; DB 21; Length 266;
Best Local Similarity 99.5%; Pred. No. 6e-91;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGGWNHHNRSSSTRSGGDLTLGLPEEEEARSPFLAPSGAGSDVFDGGLGMAAG 60
49 GAGGWNHHNRSSSTRSGGDLTLGLPEEEEARSPFLAPSGAGSDVFDGGLGMAAG 108

```

Oy	61	LOSLETHHPSPYLQRRSEDPYVLPSETGGYVAPLACSPQPEPVNPDVRRPQPSREGL	120
Dd	109	LOSLETHHPSPYLQRRSEDPYVLPSETGGYVAPLACSPQPEPVNPDVRRPQPSREGL	168
Oy	121	PAARPAAGATLEERAKTLSPGKNGVNDVFAFGAVENPEYLRPOGGAAPDHPPPAFSPA	180
Dd	169	PAARPAAGATLEERAKTLSPGKNGVNDVFAFGAVENPEYLRPOGGAAPDHPPPAFSPA	228
Oy	181	DNLVYWDQDPPERGAPPSFTFKGTTPPAENPEYLGIDVP	217
Dd	229	DNLVYWDQDPPERGAPPSFTFKGTTPPAENPEYLGIDVP	265
		RESULT 14	
		AAM51146	
ID		AAM51146 standard; Protein; 266 AA.	
XX			
AC		AAM51146;	
XX			
DT		17-JUN-2002 (first entry)	
XX			
DE		Human Her-2/neu oncoprotein phosphorylation domain.	
XX			
KW		Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;	
KM		tyrosine kinase; receptor; c-erbB2; gene therapy.	
XX			
OS		Homo sapiens.	
XX			
PN		WO200212341-A2.	
XX			
PD		14-FEB-2002.	
XX			
PF		03-AUG-2001; 2001WO-US24283.	
XX			
PR		03-AUG-2000; 2000US-0632507.	
XX			
PA		(CORI-) CORIXA CORP.	
PA		(SMIK-) SMITHKLINE BEECHAM BIOLOGICALS.	
XX			
PI		Cheever MA, Cheysen D;	
XX			
DR		WPI: 2002-241743/29.	
XX			
PT		Her-2/neu fusion protein for treating or preventing cancer by eliciting	
PT		or enhancing an immune response to the protein, has Her-2/neu	
PT		extracellular domain fused to Her-2/neu intracellular or	
PT		phosphorylation domain -	
XX			
PS		Claim 2; Fig 10; 141pp; English.	
XX			
CC		The present sequence is that of the phosphorylation domain of	
CC		human Her-2/neu (p185 glycoprotein or c-erbB2), an oncogenic	
CC		self-protein and target for anti-cancer vaccines. The Her-2/neu	
CC		gene is amplified and p185 is overexpressed in a variety of cancers,	
CC		including breast, ovarian, colon, lung and prostate cancer.	
CC		Her-2/neu (see AAM51143) is a member of the tyrosine kinase family	
CC		of receptor-like glycoproteins. It comprises an extracellular	
CC		domain with homology to the epidermal growth factor receptor	
CC		(EGFR), a highly hydrophobic transmembrane domain and a C-terminal	
CC		intracellular domain that also shows homology to EGFR. Its	
CC		overexpression correlates with a poor prognosis in breast and	
CC		ovarian cancers. The invention provides Her-2/neu fusion	
CC		comprising the fusion proteins or nucleic acid molecules. In	
CC		preferred fusion proteins, the extracellular domain of a Her-2/neu	
CC		protein is fused to a Her-2/neu intracellular domain or	
CC		phosphorylation domain (or its DeltaCD fragment). An immune	
CC		response to Her-2/neu protein is elicited or enhanced by	
CC		administering the fusion protein in the form of a vaccine, or by	
CC		transfecting cells of an animal ex vivo with a nucleic acid	
CC		encoding the fusion protein, and delivering the transfected cells	
CC		to the animal. The fusion proteins, nucleic acids, and isolated	

	CC	specific T-cells are useful for inhibiting the development of a
	CC	cancer, especially breast, ovarian, colon, lung or prostate cancer
	CC	in a patient. T cells that specifically react with a Her-2/neu
	CC	fusion protein can be used to remove tumour cells from a sample in
	CC	order to inhibit the development of cancer in a patient.
	xx	
SQ	Sequence	266 AA;
	Query Match	99.6%; Score 1177; DB 23; Length 266;
	Best Local Similarity	99.5%; Pred. No. 6e-91;
	Matches 216:	Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1	GAGGCVHHRRHSSSRSGGGDLTGLLEPSEEPAPSPPLAPSEAGSDVFPDDLGMGAAG 60
Db	49	GAGGCVHHRRHSSSRSGGGDLTGLLEPSEEPAPSPPLAPSEAGSDVFPDDLGMGAAG 108
OY	61	LQSLTHPPSPLOQRSEPTVPPLPSETDGYAAPLTCSPQPEYVNPDPVRPQPSPREGPL 120
Db	109	LQSLTHPPSPLOQRSEPTVPPLPSETDGYAAPLTCSPQPEYVNPDPVRPQPSPREGPL 168
OY	121	PAAPRAGATLERAKTLTLPKNGGVKDVAFGAVENPEYLTPQGGAQPPIPPPAFSPA 180
Db	169	PAAPRAGATLERAKTLTLPKNGGVKDVAFGAVENPEYLTPQGGAQPPIPPPAFSPA 228
OY	181	DNLYYWDDPPERGPAPSTFKGTPTAENPEYLGIDVP 217
Db	229	DNLYYWDDPPERGPAPSTFKGTPTAENPEYLGIDVP 265
	RESULT 15	
	AAE20483	
ID	AAE20483	standard; Protein: 583 AA.
XX	AC	AAE20483;
XX	DT	01-JUL-2002 (first entry)
DE	Human protein for the clone HICD_native_coding_region.	
XX	Human; Her-2/Neu protein; Immune response; gene therapy; breast cancer;	
KM	human leukocyte antigen; HLA; vaccine; malignancy; cytostatic.	
KW	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
FH	Misc-difference 581..583	
FT	/note= "Encoded by GNGRAATGCTC"	
XX	WO200214503-A2.	
PN		
XX	21-FEB-2002.	
PD		
XX	14-AUG-2001; 2001WO-US41733.	
PF		
PR	14-AUG-2000; 2000US-223152P.	
XX	28-SEP-2000; 2000US-236428P.	
PR	21-FEB-2001; 2001US-270520P.	
XX	(CORI-) CORIXA CORP.	
PA		
XX	Hand-zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos ND;	
PI	Mcneill PD, Vedvyick TS;	
XX		
DR	WPI: 2002-280758/32.	
N-PSDB:	AMD32746.	
XX		
XX	Novel isolated Her-2/Neu polypeptide composition useful for therapy,	
PT	prevention and diagnosis of cancer, preferably breast cancer -	
XX		
PS	Example 5; Page 121-122; 129pp; English.	
XX		
CC	The invention relates to an isolated Her-2/Neu polypeptide composition	
CC	effective for eliciting an immune response. The invention is useful for	

CC eliciting an immune response in a patient, where the patient is human  
CC leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.  
CC The composition is useful for the therapy and diagnosis of cancer.  
CC preferably breast cancer. In pharmaceutical compositions, e.g., vaccine  
CC and other compositions for the diagnosis, prevention and treatment of  
CC human malignancies, for stimulating and/or expanding T cells specific for  
CC Her-2/Neu polypeptide and for inhibiting the development of cancer in a  
CC patient. The invention is useful for stimulating a T cell response in a  
CC human patient, as probe or primer for nucleic acid hybridisation, to  
CC selectively form duplex molecules with complementary stretches of the  
CC entire Her-2/Neu gene or gene fragments of interest, to isolate a full  
CC length gene from a suitable library, and to direct expression of a  
CC polypeptide in appropriate host cells. The composition is useful in  
CC prophylactic or therapeutic applications and for the treatment of cancer,  
CC preferably for the immunotherapy of breast cancer and other Her-2/Neu-  
CC associated malignancies. The invention is useful in gene therapy. The  
CC present sequence is human protein for the clone H1CD\_native\_coding\_  
XX region.

SQ Sequence 583 AA;

Query Match

Best local Similarity 99.6%; Score 1177; DB 23; Length 583;

Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGMVHHRRSSSTRSGGDLTLGLEPSESEAPRSLAPSEAGSDVFDGDLGMAKG 60  
DB 364 GAGGMVHHRRSSSTRSGGDLTLGLEPSESEAPRSLAPSEAGSDVFDGDLGMAKG 423  
OY 61 LQSLPTHDPSPLQRYSEDPVPLPSEETDGYVAPLTCSPOPEYVNPQPVRRQPPSPREGPL 120  
DB 424 LQSLPTHDPSPLQRYSEDPVPLPSEETDGYVAPLTCSPOPEYVNPQPVRRQPPSPREGPL 483  
OY 121 PAARPAGATLERAKTTLSPKNGVVKDVFAGGAVENPEYLTPOGGAAPQPHPPAFSPAF 180  
DB 484 PAARPAGATLERAKTTLSPKNGVVKDVFAGGAVENPEYLTPOGGAAPQPHPPAFSPAF 543  
OY 181 DNLYYMDQDPPREGAPPTSTFKGPTAENPEYLGIDVP 217  
DB 544 DNLYYMDQDPPREGAPPTSTFKGPTAENPEYLGIDVP 580

Search completed: April 28, 2003, 13:39:47  
Job time : 18.4181 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: April 28, 2003, 13:40:24 ; Search time 7.98328 Seconds

(without alignments)  
2178.074 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182

Sequence: 1 GAGCMVHHRRSSSTRSGGCG.....STFKGPTAENPEYLGLDVP 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US06\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US07\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182	100.0	217	10	US-09-821-883-25
2	1182	100.0	397	10	US-09-821-883-27
3	1182	100.0	555	10	US-09-821-883-1
4	1182	100.0	564	10	US-09-821-883-3
5	1182	100.0	680	10	US-09-821-883-2
6	1182	100.0	697	10	US-09-821-883-4
7	1182	100.0	1179	10	US-09-821-883-29
8	1182	100.0	1255	10	US-09-769-508-2
9	1182	100.0	1255	10	US-09-811-123-9
10	1182	100.0	1255	10	US-09-811-115-3
11	1177	99.6	266	9	US-09-854-356-4
12	1177	99.6	353	9	US-09-930-125-9
13	1177	99.6	587	9	US-09-930-125-8
14	1177	99.6	589	9	US-09-930-125-10
15	1177	99.6	600	9	US-09-930-125-11
16	1177	99.6	919	9	US-09-854-356-6
17	1177	99.6	1255	9	US-09-854-356-1
18	1177	99.6	1255	9	US-09-930-125-2
19	1177	99.6	1255	9	US-09-441-411-6

20	961	81.3	1256	9	US-09-854-356-14	Sequence 14, Appl
21	954	80.7	1256	9	US-09-854-356-2	Sequence 2, Appl
22	954	80.7	1260	9	US-09-759-118	Sequence 118, App
23	711	60.2	293	9	US-10-102-806-583	Sequence 583, App
24	648.5	54.9	135	10	US-09-925-301-1232	Sequence 1232, Ap
25	164	13.9	1210	10	US-09-725-433-2	Sequence 2, Appl
26	157	13.3	1308	10	US-09-940-101-2	Sequence 2, Appl
27	155	13.1	705	9	US-09-158-722-4	Sequence 4, Appl
28	134	11.3	674	9	US-10-086-464-14	Sequence 14, Appl
29	127.5	10.8	507	9	US-10-078-547-24	Sequence 24, Appl
30	126	10.7	1274	9	US-10-020-215-2	Sequence 2, Appl
31	124	10.5	503	9	US-10-078-547-2	Sequence 2, Appl
32	121	10.2	428	9	US-09-906-514-4	Sequence 4, Appl
33	120	10.2	699	9	US-10-121-988-143	Sequence 143, App
34	118.5	10.0	731	9	US-10-086-464-8	Sequence 8, Appl
35	118	10.0	419	10	US-09-814-777A-36	Sequence 36, Appl
36	118	10.0	684	10	US-09-823-240-9	Sequence 9, Appl
37	115.5	9.8	731	9	US-10-086-464-17	Sequence 17, Appl
38	115.5	9.8	743	10	US-09-771-161A-164	Sequence 164, App
39	115.5	9.8	743	10	US-09-771-161A-254	Sequence 254, App
40	114	9.6	1464	9	US-10-060-036-159	Sequence 159, App
41	113.5	9.6	726	10	US-09-770-689A-4	Sequence 4, Appl
42	113.5	9.6	803	10	US-09-770-689A-2	Sequence 2, Appl
43	113.5	9.6	1665	10	US-09-858-664A-2	Sequence 2, Appl
44	113.5	9.6	2630	9	US-10-077-130-2	Sequence 2, Appl
45	113.5	9.6	7968	9	US-10-077-130-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1	US-09-821-883-25
Sequence 25, Application US/09821883	
Patent No. US20020061310A1	
GENERAL INFORMATION:	
APPLICANT: Laus, Rehner	
APPLICANT: Vidozic, Damir	
APPLICANT: Gradis, Thomas	
TITLE OF INVENTION: Compositions and Methods for Dendritle	
FILE REFERENCE: 7636-0022.30	
CURRENT APPLICATION NUMBER: US/09/821,883	
CURRENT FILING DATE: 2001-03-30	
PRIOR APPLICATION NUMBER: US 60/193,504	
PRIOR FILING DATE: 2000-03-30	
NUMBER OF SEQ ID NOS: 30	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 25	
LENGTH: 217	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-09-821-883-25	
Query Match	100.0%; Score 1182; DB 10; Length 217;
Best Local Similarity	100.0%; Pred. No. 5; 3e-70;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 GAGCMVHHRRSSSTRSGGDLTGLPSESEAPRSLASEGASGVFFGDLGMAAKG 60	
DB 1 GAGCMVHHRRSSSTRSGGDLTGLPSESEAPRSLASEGASGVFFGDLGMAAKG 60	
QY 61 LOSLPTDPSFLOKYSDDPYPLPSEITDGVVAPLTCSPOPEYVQPVPRQPPSPRGPL 120	
DB 61 LOSLPTDPSFLOKYSDDPYPLPSEITDGVVAPLTCSPOPEYVQPVPRQPPSPRGPL 120	
QY 121 PAARPAATLERAKTSLSPKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPAPAFSPAF 180	
DB 121 PAARPAATLERAKTSLSPKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPAPAFSPAF 180	
QY 181 DNLTYWDQDPPEKCAPSTFKGPTAENPEYLGLDVP 217	
DB 181 DNLTYWDQDPPEKCAPSTFKGPTAENPEYLGLDVP 217	

## RESULT 2

US-09-821-883-27  
; Sequence 27, Application US/09821883  
; Patent No. US20020061310A1  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Vidovic, Damir  
; APPLICANT: Gradis, Thomas  
; TITLE OF INVENTION: Compositions and Methods for Dendritic  
; FILE REFERENCE: 7636-0022.30  
; CURRENT APPLICATION NUMBER: US/09/821,883  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 60/193,504  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 397  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: NY-ESO-1C tumor antigen  
US-09-821-883-27

Query Match 100.0%; Score 1182; DB 10; Length 397;  
Best Local Similarity 100.0%; Pred. No. 1e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGMVHHRRSSSTRSGGDLTLGLEPSEEARPRSLPASEGASDVFDGDLGMAAKG 60  
DB 181 GAGGMVHHRRSSSTRSGGDLTLGLEPSEEARPRSLPASEGASDVFDGDLGMAAKG 240  
QY 61 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAPLTCSPOEYVNOQPVRRQPPSPREGPL 120  
DB 241 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAPLTCSPOEYVNOQPVRRQPPSPREGPL 300  
QY 121 PAARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
DB 301 PAARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 360  
QY 181 DNLVYMDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217  
DB 361 DNLVYMDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 397

## RESULT 3

US-09-821-883-1  
; Sequence 1, Application US/09821883  
; Patent No. US20020061310A1  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Vidovic, Damir  
; APPLICANT: Gradis, Thomas  
; TITLE OF INVENTION: Compositions and Methods for Dendritic  
; FILE REFERENCE: 7636-0022.30  
; CURRENT APPLICATION NUMBER: US/09/821,883  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 60/193,504  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 555  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HER500 construct  
US-09-821-883-1

Query Match 100.0%; Score 1182; DB 10; Length 555;  
Best Local Similarity 100.0%; Pred. No. 1.4e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGMVHHRRSSSTRSGGDLTLGLEPSEEARPRSLPASEGASDVFDGDLGMAAKG 60  
DB 330 GAGGMVHHRRSSSTRSGGDLTLGLEPSEEARPRSLPASEGASDVFDGDLGMAAKG 389  
QY 61 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAPLTCSPOEYVNOQPVRRQPPSPREGPL 120  
DB 390 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAPLTCSPOEYVNOQPVRRQPPSPREGPL 449  
QY 121 PAARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
DB 450 PAARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 509  
QY 181 DNLVYMDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217  
DB 510 DNLVYMDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 546

## RESULT 4

US-09-821-883-3  
; Sequence 3, Application US/09821883  
; Patent No. US20020061310A1  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner  
; APPLICANT: Vidovic, Damir  
; APPLICANT: Gradis, Thomas  
; TITLE OF INVENTION: Compositions and Methods for Dendritic  
; FILE REFERENCE: 7636-0022.30  
; CURRENT APPLICATION NUMBER: US/09/821,883  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: US 60/193,504  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 564  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HER500+ construct  
US-09-821-883-3

Query Match 100.0%; Score 1182; DB 10; Length 564;  
Best Local Similarity 100.0%; Pred. No. 1.4e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGMVHHRRSSSTRSGGDLTLGLEPSEEARPRSLPASEGASDVFDGDLGMAAKG 60  
DB 339 GAGGMVHHRRSSSTRSGGDLTLGLEPSEEARPRSLPASEGASDVFDGDLGMAAKG 398  
QY 61 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAPLTCSPOEYVNOQPVRRQPPSPREGPL 120  
DB 399 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAPLTCSPOEYVNOQPVRRQPPSPREGPL 458  
QY 121 PAARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
DB 459 PAARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 518  
QY 181 DNLVYMDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217  
DB 519 DNLVYMDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 555

## RESULT 5

US-09-821-883-2  
; Sequence 2, Application US/09821883  
; Patent No. US20020061310A1  
; GENERAL INFORMATION:  
; APPLICANT: Laus, Reiner



APPLICANT: Vidovic, Damir  
APPLICANT: Gradidis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
TITLE OF INVENTION: Cell-Based Immunotherapy  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821,883  
CURRENT FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/193,504  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 690  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500-HCM-CSF construct  
US-09-821-883-2

Query Match 100.0%; Score 1182; DB 10; Length 690;  
Best Local Similarity 100.0%; Pred. No. 1.8e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGVNHRHSSSTRSGGDLTGLPSEEEAPRSLAPSEAGSDVFDGDLGKAAG 60  
DB 330 GAGGVNHRHSSSTRSGGDLTGLPSEEEAPRSLAPSEAGSDVFDGDLGKAAG 389  
QY 61 LQSLPTHPSPLORYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 120  
DB 390 LQSLPTHPSPLORYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 449  
QY 121 PAAPGATLERAKTSLGKNGVYKDVAFGAVENPEYLPPOGGAPOPHPPAFSPAF 180  
DB 450 PAAPGATLERAKTSLGKNGVYKDVAFGAVENPEYLPPOGGAPOPHPPAFSPAF 509  
QY 181 DNLVYWDODPPERGAPSTFGKPTAENPEYLGIDVP 217  
DB 510 DNLVYWDODPPERGAPSTFGKPTAENPEYLGIDVP 546

RESULT 6  
US-09-821-883-4  
Sequence 4, Application US/09821883  
Patent No. US20020061310A1  
GENERAL INFORMATION:  
APPLICANT: Laus, Relner  
APPLICANT: Vidovic, Damir  
APPLICANT: Gradidis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821,883  
CURRENT FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/193,504  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 697  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HER500-HCM-CSF construct  
US-09-821-883-4

Query Match 100.0%; Score 1182; DB 10; Length 697;  
Best Local Similarity 100.0%; Pred. No. 1.8e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGVNHRHSSSTRSGGDLTGLPSEEEAPRSLAPSEAGSDVFDGDLGKAAG 60  
DB 330 GAGGVNHRHSSSTRSGGDLTGLPSEEEAPRSLAPSEAGSDVFDGDLGKAAG 398

QY 61 LQSLPTHPSPLORYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 120  
DB 390 LQSLPTHPSPLORYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 458  
QY 121 PAAPGATLERAKTSLGKNGVYKDVAFGAVENPEYLPPOGGAPOPHPPAFSPAF 180  
DB 459 PAAPGATLERAKTSLGKNGVYKDVAFGAVENPEYLPPOGGAPOPHPPAFSPAF 518  
QY 181 DNLVYWDODPPERGAPSTFGKPTAENPEYLGIDVP 217  
DB 519 DNLVYWDODPPERGAPSTFGKPTAENPEYLGIDVP 555

RESULT 7  
US-09-821-883-29  
Sequence 29, Application US/09821883  
Patent No. US20020061310A1  
GENERAL INFORMATION:  
APPLICANT: Laus, Relner  
APPLICANT: Vidovic, Damir  
APPLICANT: Gradidis, Thomas  
TITLE OF INVENTION: Compositions and Methods for Dendritic  
FILE REFERENCE: 7636-0022.30  
CURRENT APPLICATION NUMBER: US/09/821,883  
CURRENT FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/193,504  
PRIOR FILING DATE: 2000-03-30  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 29  
LENGTH: 1179  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: SART-3-IC  
US-09-821-883-29

Query Match 100.0%; Score 1182; DB 10; Length 1179;  
Best Local Similarity 100.0%; Pred. No. 3.1e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGVNHRHSSSTRSGGDLTGLPSEEEAPRSLAPSEAGSDVFDGDLGKAAG 60  
DB 963 GAGGVNHRHSSSTRSGGDLTGLPSEEEAPRSLAPSEAGSDVFDGDLGKAAG 1022  
QY 61 LQSLPTHPSPLORYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 120  
DB 1023 LQSLPTHPSPLORYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 1082  
QY 121 PAAPGATLERAKTSLGKNGVYKDVAFGAVENPEYLPPOGGAPOPHPPAFSPAF 180  
DB 1083 PAAPGATLERAKTSLGKNGVYKDVAFGAVENPEYLPPOGGAPOPHPPAFSPAF 1142  
QY 181 DNLVYWDODPPERGAPSTFGKPTAENPEYLGIDVP 217  
DB 1143 DNLVYWDODPPERGAPSTFGKPTAENPEYLGIDVP 1179

RESULT 8  
US-09-769-508-2  
Sequence 2, Application US/09769508  
Patent No. US20020155527A1  
GENERAL INFORMATION:  
APPLICANT: STUART, SUSAN G.  
APPLICANT: MONAHAN, JOHN J.  
APPLICANT: LANGTON, BEATRICE CLAUDIA  
APPLICANT: HANCOCK, MIRIAM E.C.  
APPLICANT: CHAO, LORRINE A.  
APPLICANT: BLUFORD, PETER A.  
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75  
FILE REFERENCE: BBIO-111-C1  
CURRENT APPLICATION NUMBER: US/09/769,508

;; CURRENT FILING DATE: 2001-01-26  
;; NUMBER OF SEQ ID NOS: 2  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 2  
;; LENGTH: 1255  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-769-508-2

Query Match 100.0%; Score 1182; DB 9; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 3.3e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAVHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGAGSDVFPDGLGMGAAG 60  
DB 1038 GAGGAVHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGAGSDVFPDGLGMGAAG 1097  
QY 61 LQSLPTHDPSPLOKRSYSDPTVPLPSETDGYVAPLTCSPOPEYVNPDPVRPQPSPRGGL 120  
DB 1098 LQSLPTHDPSPLOKRSYSDPTVPLPSETDGYVAPLTCSPOPEYVNPDPVRPQPSPRGGL 1157  
QY 121 PAARPAGATLERAKTISPCKNGYVKDVFARFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
DB 1158 PAARPAGATLERAKTISPCKNGYVKDVFARFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 1217  
QY 181 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217  
DB 1218 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 9  
US-09-811-123-9  
; Sequence 9, Application US/09811123  
; Patent No. US20020001587A1  
; GENERAL INFORMATION:  
; APPLICANT: Sharon Erickson  
; APPLICANT: Ralph Schwall  
; APPLICANT: Mark Sliwowski  
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EBB  
; FILE REFERENCE: GENENT. 073A2  
; CURRENT APPLICATION NUMBER: US/09/811,123  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/238,327  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 09/602,530  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-811-123-9

Query Match 100.0%; Score 1182; DB 10; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 3.3e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAVHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGAGSDVFPDGLGMGAAG 60  
DB 1038 GAGGAVHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGAGSDVFPDGLGMGAAG 1097  
QY 61 LQSLPTHDPSPLOKRSYSDPTVPLPSETDGYVAPLTCSPOPEYVNPDPVRPQPSPRGGL 120  
DB 1098 LQSLPTHDPSPLOKRSYSDPTVPLPSETDGYVAPLTCSPOPEYVNPDPVRPQPSPRGGL 1157  
QY 121 PAARPAGATLERAKTISPCKNGYVKDVFARFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
DB 1158 PAARPAGATLERAKTISPCKNGYVKDVFARFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 1217  
QY 181 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217

DB 1218 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 10  
US-09-811-115-3  
; Sequence 3, Application US/09811115  
; Patent No. US20020035736A1  
; GENERAL INFORMATION:  
; APPLICANT: Erickson, Sharon  
; APPLICANT: Schwall, Ralph  
; APPLICANT: King, Kathleen  
; TITLE OF INVENTION: HER-2/TRANSGENIC NON-HUMAN TUMOR MODEL  
; FILE REFERENCE: GENENT. 034A  
; CURRENT APPLICATION NUMBER: US/09/811,115  
; CURRENT FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/189,844  
; PRIOR FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1255  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-811-115-3

Query Match 100.0%; Score 1182; DB 10; Length 1255;  
Best Local Similarity 100.0%; Pred. No. 3.3e-69;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGAVHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGAGSDVFPDGLGMGAAG 60  
DB 1038 GAGGAVHHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGAGSDVFPDGLGMGAAG 1097  
QY 61 LQSLPTHDPSPLOKRSYSDPTVPLPSETDGYVAPLTCSPOPEYVNPDPVRPQPSPRGGL 120  
DB 1098 LQSLPTHDPSPLOKRSYSDPTVPLPSETDGYVAPLTCSPOPEYVNPDPVRPQPSPRGGL 1157  
QY 121 PAARPAGATLERAKTISPCKNGYVKDVFARFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 180  
DB 1158 PAARPAGATLERAKTISPCKNGYVKDVFARFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 1217  
QY 181 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217  
DB 1218 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 11  
US-09-854-356-4  
; Sequence 4, Application US/09854356  
; Patent No. US20020177567A1  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Gheysen, Dirk  
; APPLICANT: Corixa Corporation  
; APPLICANT: SmithKline Beecham Biologicals S. A.  
; TITLE OF INVENTION: HER-2/neu Fusion Proteins  
; FILE REFERENCE: 014058-009810PC  
; CURRENT APPLICATION NUMBER: US/09/854,356  
; CURRENT FILING DATE: 2001-05-09  
; PRIOR APPLICATION NUMBER: US 09/493,480  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,976  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu  
US-09-854-356-4

Query Match 99.6%; Score 1177; DB 9; Length 266;  
Best Local Similarity 99.5%; Pred. No. 1.4e-69;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 60  
DB 49 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 108  
OY 61 LOSLTHDPSPLQRYSEDPYVLPSETDGYVAPLTCSPQPEYVNOPDVAPPPSPREGPL 120  
DB 109 LOSLTHDPSPLQRYSEDPYVLPSETDGYVAPLTCSPQPEYVNOPDVAPPPSPREGPL 168  
OY 121 PAARAGATLERAKTLPSPGKNGVVDVAFAGAVENPEYLPFGGAAPQHPHPPAFSPAF 180  
DB 169 PAARAGATLERAKTLPSPGKNGVVDVAFAGAVENPEYLPFGGAAPQHPHPPAFSPAF 228  
OY 181 DNLVYWDODPPERGAPSPSTFKCTPTAENPEYLGLDVP 217  
DB 229 DNLVYWDODPPERGAPSPSTFKCTPTAENPEYLGLDVP 265

## RESULT 12

US-09-930-125-9  
Sequence 9, Application US/09930125  
Publication No. US20020193329A1  
GENERAL INFORMATION:  
APPLICANT: Hand-Zimmerman, Susan  
APPLICANT: Cheever, Martin A.  
APPLICANT: Foy, Teresa M.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Katos, Michael D.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Vegdick, Thomas S.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
FILE REFERENCE: 210121.544  
CURRENT APPLICATION NUMBER: US/09/930.125  
CURRENT FILING DATE: 2001-08-14  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 9  
LENGTH: 583  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-930-125-9

Query Match 99.6%; Score 1177; DB 9; Length 583;  
Best Local Similarity 99.5%; Pred. No. 3.1e-69;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 60  
DB 364 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 423  
OY 61 LOSLTHDPSPLQRYSEDPYVLPSETDGYVAPLTCSPQPEYVNOPDVAPPPSPREGPL 120  
DB 424 LOSLTHDPSPLQRYSEDPYVLPSETDGYVAPLTCSPQPEYVNOPDVAPPPSPREGPL 483  
OY 121 PAARAGATLERAKTLPSPGKNGVVDVAFAGAVENPEYLPFGGAAPQHPHPPAFSPAF 180  
DB 484 PAARAGATLERAKTLPSPGKNGVVDVAFAGAVENPEYLPFGGAAPQHPHPPAFSPAF 543  
OY 181 DNLVYWDODPPERGAPSPSTFKCTPTAENPEYLGLDVP 217  
DB 544 DNLVYWDODPPERGAPSPSTFKCTPTAENPEYLGLDVP 580

RESULT 13  
US-09-930-125-8  
Sequence 8, Application US/09930125  
Publication No. US20020193329A1  
GENERAL INFORMATION:  
APPLICANT: Hand-Zimmerman, Susan

APPLICANT: Cheever, Martin A.  
APPLICANT: Foy, Teresa M.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Katos, Michael D.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Vegdick, Thomas S.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
FILE REFERENCE: 210121.544  
CURRENT APPLICATION NUMBER: US/09/930.125  
CURRENT FILING DATE: 2001-08-14  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 587  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-930-125-8

Query Match 99.6%; Score 1177; DB 9; Length 587;  
Best Local Similarity 99.5%; Pred. No. 3.2e-69;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 60  
DB 364 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 423  
OY 61 LOSLTHDPSPLQRYSEDPYVLPSETDGYVAPLTCSPQPEYVNOPDVAPPPSPREGPL 120  
DB 424 LOSLTHDPSPLQRYSEDPYVLPSETDGYVAPLTCSPQPEYVNOPDVAPPPSPREGPL 483  
OY 121 PAARAGATLERAKTLPSPGKNGVVDVAFAGAVENPEYLPFGGAAPQHPHPPAFSPAF 180  
DB 484 PAARAGATLERAKTLPSPGKNGVVDVAFAGAVENPEYLPFGGAAPQHPHPPAFSPAF 543  
OY 181 DNLVYWDODPPERGAPSPSTFKCTPTAENPEYLGLDVP 217  
DB 544 DNLVYWDODPPERGAPSPSTFKCTPTAENPEYLGLDVP 580

## RESULT 14

US-09-930-125-10  
Sequence 10, Application US/09930125  
Publication No. US20020193329A1  
GENERAL INFORMATION:  
APPLICANT: Hand-Zimmerman, Susan  
APPLICANT: Cheever, Martin A.  
APPLICANT: Foy, Teresa M.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Katos, Michael D.  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Vegdick, Thomas S.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS  
TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES  
FILE REFERENCE: 210121.544  
CURRENT APPLICATION NUMBER: US/09/930.125  
CURRENT FILING DATE: 2001-08-14  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 589  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-930-125-10

Query Match 99.6%; Score 1177; DB 9; Length 589;  
Best Local Similarity 99.5%; Pred. No. 3.2e-69;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 60  
DB 372 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPRSPPLAPSGAGSDVFDGDLGGAAG 431



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 28, 2003, 13:39:54 ; Search time 6.3503 Seconds  
(without alignments)  
1005.423 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182  
Sequence: 1 GAGGCVHHRSSSTRSGGCG.....STFKGPTAENPEYLGIDVP 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : Issued Patents-AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCUTOS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1177	99.6	580	1 US-08-414-417B-69	Sequence 69, Appl
2	1177	99.6	580	2 US-08-486-348A-69	Sequence 69, Appl
3	1177	99.6	580	3 US-08-468-545B-69	Sequence 69, Appl
4	1177	99.6	580	4 US-08-466-680B-69	Sequence 68, Appl
5	1177	99.6	1255	1 US-08-467-083-68	Sequence 68, Appl
6	1177	99.6	1255	1 US-08-414-417B-68	Sequence 68, Appl
7	1177	99.6	1255	2 US-08-486-348A-68	Sequence 68, Appl
8	1177	99.6	1255	2 US-08-625-101-2	Sequence 2, Appl
9	1177	99.6	1255	2 US-08-468-545B-68	Sequence 68, Appl
10	1177	99.6	1255	2 US-08-356-766-2	Sequence 2, Appl
11	1177	99.6	1255	3 US-08-466-680B-68	Sequence 68, Appl
12	1157	97.9	1255	3 US-08-484-438-8	Sequence 8, Appl
13	164	13.9	1210	2 US-08-484-438-7	Sequence 7, Appl
14	164	13.9	1210	2 US-08-475-035-4	Sequence 4, Appl
15	157	13.3	541	2 US-08-484-438-6	Sequence 4, Appl
16	157	13.3	1308	2 US-08-484-438-2	Sequence 2, Appl
17	155	13.1	705	2 US-08-456-647B-4	Sequence 4, Appl
18	155	13.1	705	2 US-08-237-401A-4	Sequence 4, Appl
19	126.5	10.7	1321	2 US-08-317-310A-64	Sequence 64, Appl
20	126	10.6	1274	4 US-09-095-443-2	Sequence 2, Appl
21	125.5	10.6	314	2 US-08-525-742-6	Sequence 6, Appl
22	116	9.8	1248	2 US-09-080-897-2	Sequence 2, Appl
23	116	9.8	1248	4 US-09-323-735-2	Sequence 2, Appl
24	115	9.7	124	4 US-08-925-237-2	Sequence 2, Appl
25	114	9.6	1057	3 US-08-931-820-1	Sequence 1, Appl
26	114	9.6	1341	3 US-08-963-825-18	Sequence 18, Appl
27	114	9.6	1341	4 US-09-500-811-18	Sequence 18, Appl

28	114	9.6	1341	4 US-09-570-573-18	Sequence 18, Appl
29	114	9.6	1341	4 US-09-548-608-18	Sequence 18, Appl
30	114	9.6	1461	4 US-09-585-887-9	Sequence 9, Appl
31	114	9.6	1461	4 US-09-289-578-9	Sequence 9, Appl
32	113.5	9.6	2972	4 US-09-579-181-2	Sequence 2, Appl
33	113.5	9.6	3118	4 US-09-579-181-1	Sequence 1, Appl
34	111	9.4	627	4 US-08-487-596-6	Sequence 6, Appl
35	111	9.4	696	3 US-08-906-865-4	Sequence 4, Appl
36	111	9.4	696	4 US-09-129-668-4	Sequence 4, Appl
37	110	9.3	1185	4 US-09-041-886-23	Sequence 23, Appl
38	110	9.3	1315	3 US-08-899-595-3	Sequence 3, Appl
39	110	9.3	1317	3 US-09-083-521-7	Sequence 7, Appl
40	109.5	9.3	1706	2 US-08-459-568-2	Sequence 2, Appl
41	109.5	9.3	1706	2 US-08-399-411-2	Sequence 2, Appl
42	109.5	9.3	1706	3 US-08-516-859A-2	Sequence 2, Appl
43	109.5	9.3	1706	4 US-09-586-472-2	Sequence 2, Appl
44	109.5	9.3	1706	4 US-09-528-706-2	Sequence 2, Appl
45	108.5	9.2	214	1 US-08-217-327-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-414-417B-69  
Sequence 69, Application US/0841417B  
Patent No. 5801005  
GENERAL INFORMATION:  
APPLICANT: Cheever, Martin A.  
APPLICANT: Disla, Mary L.  
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/414,417B  
FILING DATE: 31-MAR-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Sharkey, Richard G.  
REGISTRATION NUMBER: 32,629  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 580 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-414-417B-69

Query Match 99.6%; Score 1177; DB 1; Length 580;  
Best Local Similarity 99.5%; Pred. No. 3.3e-96;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGCVHHRSSSTRSGGDLTLGLEPSEEPAPSPPLASRGAGSDVDPGLJMGAAKG 60  
DB 363 GAGGCVHHRSSSTRSGGDLTLGLEPSEEPAPSPPLASRGAGSDVDPGLJMGAAKG 422

QY 61 LOSLPTHDPSPLORXSDDPTVPLPSETDGVAPLTCSPQPEYVQNPVRRPQPSPREGCL 120  
DB 423 LOSLPTHDPSPLORXSDDPTVPLPSETDGVAPLTCSPQPEYVQNPVRRPQPSPREGCL 482  
QY 121 PAARPGATLERAKTISPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPHPPAFSPAF 180  
DB 483 PAARPGATLERAKTISPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPHPPAFSPAF 542  
QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
DB 543 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 579

RESULT 2  
US-08-486-348A-69  
; Sequence 69, Application US/08486348A  
; Patent No. 5846538  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,348A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION/DOCKET NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 580 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-486-348A-69

Query Match 99.6%; Score 1177; DB 2; Length 580;  
Best Local Similarity 99.5%; Pred. No. 3.3e-96;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGMVHRRSSSTRSGGDLTGLPSEBEARSPPLASSEGASDVFPGLGMAKAG 60  
DB 363 GAGGMVHRRSSSTRSGGDLTGLPSEBEARSPPLASSEGASDVFPGLGMAKAG 422  
QY 61 LOSLPTHDPSPLORXSDDPTVPLPSETDGVAPLTCSPQPEYVQNPVRRPQPSPREGCL 120  
DB 423 LOSLPTHDPSPLORXSDDPTVPLPSETDGVAPLTCSPQPEYVQNPVRRPQPSPREGCL 482  
QY 121 PAARPGATLERAKTISPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPHPPAFSPAF 180  
DB 483 PAARPGATLERAKTISPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPHPPAFSPAF 542  
QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
DB 543 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 579

DB 543 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 579

RESULT 3  
US-08-468-545B-69  
; Sequence 69, Application US/08468545B  
; Patent No. 5876712  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.  
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,545B  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sharkey, Richard G.  
; REGISTRATION/DOCKET NUMBER: 32,629  
; REFERENCE/DOCKET NUMBER: 920010.448C5  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 580 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-468-545B-69

Query Match 99.6%; Score 1177; DB 2; Length 580;  
Best Local Similarity 99.5%; Pred. No. 3.3e-96;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGMVHRRSSSTRSGGDLTGLPSEBEARSPPLASSEGASDVFPGLGMAKAG 60  
DB 363 GAGGMVHRRSSSTRSGGDLTGLPSEBEARSPPLASSEGASDVFPGLGMAKAG 422  
QY 61 LOSLPTHDPSPLORXSDDPTVPLPSETDGVAPLTCSPQPEYVQNPVRRPQPSPREGCL 120  
DB 423 LOSLPTHDPSPLORXSDDPTVPLPSETDGVAPLTCSPQPEYVQNPVRRPQPSPREGCL 482  
QY 121 PAARPGATLERAKTISPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPHPPAFSPAF 180  
DB 483 PAARPGATLERAKTISPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPHPPAFSPAF 542  
QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
DB 543 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 579

RESULT 4  
US-08-466-680B-69  
; Sequence 69, Application US/08466680B  
; Patent No. 6075122  
; GENERAL INFORMATION:  
; APPLICANT: Cheever, Martin A.  
; APPLICANT: Disis, Mary L.

;; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
;; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
;; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
;; NUMBER OF SEQUENCES: 69  
;; CORRESPONDENCE ADDRESSES:  
;; ADDRESSEE: Seed and Berry LLP  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: US  
;; ZIP: 98104-7092  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/466,680B  
;; FILING DATE: 06-JUN-1995  
;; CLASSIFICATION: 424  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sharkey, Richard G.  
;; REGISTRATION NUMBER: 32,629  
;; REFERENCE/DOCKET NUMBER: 920010.448C4  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031  
;;  
;; INFORMATION FOR SEQ ID NO: 69:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 580 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;;  
US-08-466-680B-69

Query Match 99.6%; Score 1177; DB 3; Length 580;  
Best Local Similarity 99.5%; Pred. No. 3,3e-96;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGVNHHRRSSSTRSGGDLTGLPESEEPAPSPAPSGAGSDVFDGGLGKAKG 60  
DB 363 GAGGVNHHRRSSSTRSGGDLTGLPESEEPAPSPAPSGAGSDVFDGGLGKAKG 422  
QY 61 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 120  
DB 423 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 482  
QY 121 PAAPAGATLERAKTLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQPHPPAFSPA 180  
DB 483 PAAPAGATLERAKTLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQPHPPAFSPA 542  
QY 181 DNLVYWDODPPERGAPSTFGKTPAENPEYLGIDVP 217  
DB 543 DNLVYWDODPPERGAPSTFGKTPAENPEYLGIDVP 579

## RESULT 5

US-08-467-083-68  
;; Sequence 68, Application US/08467083  
;; Patent No. 5726023  
;; GENERAL INFORMATION:  
;; APPLICANT: Cheever, Martin A.  
;; APPLICANT: Disis, Mary L.  
;; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
;; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
;; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
;; NUMBER OF SEQUENCES: 68  
;; CORRESPONDENCE ADDRESSES:  
;; ADDRESSEE: Seed and Berry  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: US

;; ZIP: 98104-7092  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/467,083  
;; FILING DATE: 06-JUN-1995  
;; CLASSIFICATION: 424  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/414,417  
;; FILING DATE: 06-JUN-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sharkey, Richard G.  
;; REGISTRATION NUMBER: 32,629  
;; REFERENCE/DOCKET NUMBER: 920010.448C2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (206) 622-4900  
;; TELEFAX: (206) 682-6031  
;; TELEX: 3723836 SEEDANBERY  
;;  
;; INFORMATION FOR SEQ ID NO: 68:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1255 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;;  
US-08-467-083-68

Query Match 99.6%; Score 1177; DB 1; Length 1255;  
Best Local Similarity 99.5%; Pred. No. 8,5e-96;  
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGVNHHRRSSSTRSGGDLTGLPESEEPAPSPAPSGAGSDVFDGGLGKAKG 60  
DB 1038 GAGGVNHHRRSSSTRSGGDLTGLPESEEPAPSPAPSGAGSDVFDGGLGKAKG 1097  
QY 61 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 120  
DB 1098 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 1157  
QY 121 PAAPAGATLERAKTLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQPHPPAFSPA 180  
DB 1158 PAAPAGATLERAKTLSGKNGVYKDVAFGAVENPEYLTPOGGAAPQPHPPAFSPA 1217  
QY 181 DNLVYWDODPPERGAPSTFGKTPAENPEYLGIDVP 217  
DB 1218 DNLVYWDODPPERGAPSTFGKTPAENPEYLGIDVP 1254

## RESULT 6

US-08-414-417B-68  
;; Sequence 68, Application US/08414417B  
;; Patent No. 3601005  
;; GENERAL INFORMATION:  
;; APPLICANT: Cheever, Martin A.  
;; APPLICANT: Disis, Mary L.  
;; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN  
;; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE  
;; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED  
;; NUMBER OF SEQUENCES: 69  
;; CORRESPONDENCE ADDRESSES:  
;; ADDRESSEE: Seed and Berry LLP  
;; STREET: 6300 Columbia Center, 701 Fifth Avenue  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: US  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:

```

: APPLICATION NUMBER: US/08/414,417B
: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-4900
: TELEFAX: (206) 622-4900
: INFORMATION FOR SEQ ID NO: 68:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
US-08-414-417B-68

Query Match          99.6%; Score 1177; DB 1; Length 1255;
Best Local Similarity 99.5%; Pred. No. 8.5e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGVVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMCAKG 60
DB 1038 GAGGVVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMCAKG 1097
QY 61 LQSLPTHPDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVRPQPPSPREGPL 120
DB 1098 LQSLPTHPDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVRPQPPSPREGPL 1157
QY 121 PAARPAGATLERAKTISPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 180
DB 1158 PAARPAGATLERAKTISPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 1217
QY 181 DNLVYMDQDPERGAPPSTFKGTPTAENPEYLGIDVP 217
DB 1218 DNLVYMDQDPERGAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 7
US-08-486-348A-68
: Sequence 68, Application US/08486348A
: Patent No. 5846538
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Disis, Mary L.
: TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
: TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
: TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
: NUMBER OF SEQUENCES: 69
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seed and Berry LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/486,348A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C6
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 68:

```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
US-08-486-348A-68

Query Match          99.6%; Score 1177; DB 2; Length 1255;
Best Local Similarity 99.5%; Pred. No. 8.5e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGVVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMCAKG 60
DB 1038 GAGGVVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGAGSDVFDGDLGMCAKG 1097
QY 61 LQSLPTHPDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVRPQPPSPREGPL 120
DB 1098 LQSLPTHPDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVRPQPPSPREGPL 1157
QY 121 PAARPAGATLERAKTISPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 180
DB 1158 PAARPAGATLERAKTISPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 1217
QY 181 DNLVYMDQDPERGAPPSTFKGTPTAENPEYLGIDVP 217
DB 1218 DNLVYMDQDPERGAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 8
US-08-625-101-2
: Sequence 2, Application US/08625101
: Patent No. 5869445
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Disis, Mary L.
: TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
: TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
: TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
: TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED AND BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/625,101
: FILING DATE: 01-APR-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C7
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-625-101-2

Query Match          99.6%; Score 1177; DB 2; Length 1255;
Best Local Similarity 99.5%; Pred. No. 8.5e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```





## RESULT 11

US-08-466-680B-68

Sequence 68, Application US/08466680B

Patent No. 6075122

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Dals, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,680B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ. ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-466-680B-68

Query Match 99.6%; Score 1177; DB 3; Length 1255;

Best Local Similarity 99.5%; Pred. No. 8.5e-96;

Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	GAGGWHRRSSSTRSGGDLTLGLEPSEEA	PRSP	PLAS	SEGAGSDV	FDGDLG	MAK	60
DB	1038	GAGGWHRRSSSTRSGGDLTLGLEPSEEA <td>PRSP <td>PLAS <th>SEGAGSDV</th> <th>FDGDLG</th> <th>MAK</th> <th>1097</th> </td></td>	PRSP <td>PLAS <th>SEGAGSDV</th> <th>FDGDLG</th> <th>MAK</th> <th>1097</th> </td>	PLAS <th>SEGAGSDV</th> <th>FDGDLG</th> <th>MAK</th> <th>1097</th>	SEGAGSDV	FDGDLG	MAK	1097

QY	61	LOSPLTHDPSPLORYSEDPVPLPSETDGYVA	PLTCS	POPEYV	NQDP	VRPQ	PPSP	REGPL	120
DB	1098	LOSPLTHDPSPLORYSEDPVPLPSETDGYVA	PLTCS	POPEYV	NQDP	VRPQ	PPSP	REGPL	1157

QY	121	PAARPGATLERAKTLSPKNGVVKDVFAR	GAV	VENPEY	LTPOGGA	POPHPP	PAFSP	180
DB	1158	PAARPGATLERAKTLSPKNGVVKDVFAR	GAV	VENPEY	LTPOGGA	POPHPP	PAFSP	1217

QY	181	DNLTYMDQDPERGAPSTFKGTPTAENPEY	GLDVP	217
DB	1218	DNLTYMDQDPERGAPSTFKGTPTAENPEY	GLDVP	1254

## RESULT 12

US-08-484-438-8

Sequence 8, Application US/08484438

Patent No. 5811098

Patent No. 5811098 5780031

GENERAL INFORMATION:

APPLICANT: PLOWMAN, Gregory D.

APPLICANT: CULOUSCOU, Jean-Michel

APPLICANT: SHOYAB, Mohammed

APPLICANT: STEGALL, Clay B.

APPLICANT:	HELLSTR M, INGEGERD
APPLICANT:	HELLSTR M, KARI E.
TITLE OF INVENTION:	HER4 HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES:	42
CORRESPONDENCE ADDRESS:	
ADDRESSEE:	Pennie & Edmonds
STREET:	1155 Avenue of the Americas
CITY:	New York
STATE:	New York
COUNTRY:	U.S.A.
ZIP:	10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,438

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/323,442

FILING DATE: 14-OCT-1994

APPLICATION NUMBER: US 08/150,704

FILING DATE: 10-NOV-1993

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/981,165

FILING DATE: 24-NOV-1992

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: MISTOCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 5624-230

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-8090

TELEFAX: (212) 869-8864/9741

INFORMATION FOR SEQ. ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-484-438-8

Query Match 97.9%; Score 1157; DB 2; Length 1255;

Best Local Similarity 99.1%; Pred. No. 5e-94;

Matches 216; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY	1	GAGGWHRRSSSTRSGGDLTLGLEPSEEA	PRSP	PLAS	SEGAGSDV	FDGDLG	MAK	60
DB	1038	GAGGWHRRSSSTRSGGDLTLGLEPSEEA	PRSP	PLAS	SEGAGSDV	FDGDLG	MAK	1097

QY	61	LOSPLTHDPSPLORYSEDPVPLPSETDGYVA	PLTCS	POPEYV	NQDP	VRPQ	PPSP	REGPL	120
DB	1098	LOSPLTHDPSPLORYSEDPVPLPSETDGYVA	PLTCS	POPEYV	NQDP	VRPQ	PPSP	REGPL	1157

QY	121	PAARPGATLERAKTLSPKNGVVKDVFAR	GAV	VENPEY	LTPOGGA	POPHPP	PAFSP	180
DB	1158	PAARPGATLERAKTLSPKNGVVKDVFAR	GAV	VENPEY	LTPOGGA	POPHPP	PAFSP	1217

QY	181	DNLTYMDQDPERGAPSTFKGTPTAENPEY	GLDVP	217
DB	1218	DNLTYMDQDPERGAPSTFKGTPTAENPEY	GLDVP	1254

## RESULT 13

US-08-484-438-7

Sequence 7, Application US/08484438

Patent No. 5811098

Patent No. 5811098 5780031

GENERAL INFORMATION:  
APPLICANT: PLOWMAN, Gregory D.  
APPLICANT: CULOUSCOU, Jean-Michel  
APPLICANT: SHOYAB, Mohammed  
APPLICANT: SIEGALL, Clay B.  
APPLICANT: HELISTR M, Ingegerd  
APPLICANT: HELISTR M, Karl E.  
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Penille & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,438  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/323,442  
FILING DATE: 14-OCT-1994  
APPLICATION NUMBER: US 08/150,704  
FILING DATE: 10-NOV-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/981,165  
FILING DATE: 24-NOV-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1210 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-438-7

Query Match 13.9%; Score 164; DB 2; Length 1210;  
Best Local Similarity 30.5%; Pred. No. 2.5e-06;  
Matches 62; Conservative 24; Mismatches 69; Indels 48; Gaps 11;

OY 29 SEEARPSRLAPSEAGSDVFDGDLGMAKGLSLPTHDSPLDORYSEDPVPLPSET- 87  
DB 1025 SSPSTSRTPLLSLSSATSN--NSTVACIDRNGLOSCPIKEDSFLORYSSDPTGALTEDI 1082  
OY 88 -DGYVAPLTCSPOPEYVNOQVRRPQPSRREGPLPAARPGATLERRAKTLSPGKNGVYKD 146  
DB 1083 DDTFL-----PVPEYING-SVKKRPAGSVQNPVYHNQPLNP-----APSRPHQD 1127  
OY 147 VFAGGAVENPEYL-TPQGGAAPQPHPPAPSPAFDNLTYMDQ-----DP----- 190  
DB 1128 --PHSTANGNPEYLTWQ-----PTCVNSTFDSAPHAHQKSHQISLNDPYYQDDE 1176  
OY 191 -PERGAPSTFKGTPTAENPEYL 212  
DB 1177 FPKKAPNGCIFKGS-TAENAEYL 1198

RESULT 14  
US-08-475-035-4  
Sequence 4, Application US/08475035  
Patent No. 5985553  
GENERAL INFORMATION:  
APPLICANT: KING, C. R.  
APPLICANT: KRAUS, MATTHIAS H.  
APPLICANT: AARONSON, STUART A.  
TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NEEDLE & ROSENBERG, P.C.  
STREET: Suite 1200, 127 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,035  
FILING DATE: 7 Jun 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pettyman, David G.  
REGISTRATION NUMBER: 33,438  
REFERENCE/DOCKET NUMBER: 1414,656  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1210 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-475-035-4

Query Match 13.9%; Score 164; DB 2; Length 1210;  
Best Local Similarity 30.5%; Pred. No. 2.5e-06;  
Matches 62; Conservative 24; Mismatches 69; Indels 48; Gaps 11;

OY 29 SEEARPSRLAPSEAGSDVFDGDLGMAKGLSLPTHDSPLDORYSEDPVPLPSET- 87  
DB 1025 SSPSTSRTPLLSLSSATSN--NSTVACIDRNGLOSCPIKEDSFLORYSSDPTGALTEDI 1082  
OY 88 -DGYVAPLTCSPOPEYVNOQVRRPQPSRREGPLPAARPGATLERRAKTLSPGKNGVYKD 146  
DB 1083 DDTFL-----PVPEYING-SVKKRPAGSVQNPVYHNQPLNP-----APSRPHQD 1127  
OY 147 VFAGGAVENPEYL-TPQGGAAPQPHPPAPSPAFDNLTYMDQ-----DP----- 190  
DB 1128 --PHSTANGNPEYLTWQ-----PTCVNSTFDSAPHAHQKSHQISLNDPYYQDDE 1176  
OY 191 -PERGAPSTFKGTPTAENPEYL 212  
DB 1177 FPKKAPNGCIFKGS-TAENAEYL 1198

RESULT 15  
US-08-484-438-6  
Sequence 6, Application US/08484438  
Patent No. 5811098  
Patent No. 5811098 5780031  
GENERAL INFORMATION:  
APPLICANT: PLOWMAN, Gregory D.  
APPLICANT: CULOUSCOU, Jean-Michel  
APPLICANT: SHOYAB, Mohammed  
APPLICANT: SIEGALL, Clay B.

APPLICANT: Hellstr m, Ingegerd  
APPLICANT: Hellstr m, Karl E.  
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,438  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/323,442  
FILING DATE: 14-OCT-1994  
APPLICATION NUMBER: US 08/150,704  
FILING DATE: 10-NOV-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/981,165  
FILING DATE: 24-NOV-1992  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 5624-230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNITE  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 541 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-438-6

Query Match 13.3%; Score 157; DB 2; length 541;  
Best local Similarity 24.2%; Pred. No. 3.9e-06;  
Matches 62; Conservative 23; Mismatches 75; Indels 96; Gaps 10;

QY 16 RSGGGDTTGLLE-PSEERAPRSPLAP-SEGAGSDVFDGDLGMAKGLQSLPTHPSPPLQ 73  
DB 300 RDGFGFAEGVSVYRAPRTSTIEAPVAGATAEIFDDSCNGTLRKPVAPHVQEDSSTQ 359  
QY 74 RYSEDPVTPPLPS-----ETDGYVAPLTCSPQPEYVNOQDVNRQPPSPREGPLPARPA 126  
DB 360 RYSADPTVFAPERSPRGELDEEGYMPMRDKPQOXYLNVE-----ENPFVSR-- 408  
QY 127 GATLERAKTLPSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPPPA----- 175  
DB 409 -----KNGDLQ-----ALDNPETHNASNG-----PPKADEYVNEPLYL 442  
QY 176 -----FSPAFDNLVYWDODPPERGA--PPSTFKGTP----- 205  
DB 443 NTFANTLGKAEYLKNNILNSPEKAKKAFDNDPWNHSLPPRSTLQHPDYLOEYSTKYFYK 502  
QY 206 -----AENPEYL 212  
DB 503 QNGRIRPIVANPEYL 518

Search completed: April 28, 2003, 13:43:20  
Job time: 8.35033 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:49 ; Search time 14.515 Seconds  
(without alignments)  
3080.404 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182

Sequence: 1 GAGCGVHHRHRSSTRSGCG.....STFKGTPAENPEYLGLDVP 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1026.5	86.8	1259	6 018735	018735 canis famli
2	961	81.3	1367	11 08R2X1	08R2X1 mus musculu
3	906	76.6	165	4 014256	014256 homo sapien
4	894.5	75.7	412	4 08WYV0	08WYV0 homo sapien
5	144.5	14.0	1137	13 09W6F6	09W6F6 gallus galli
6	144.5	12.2	420	5 08SZ47	08SZ47 drosophila
7	144.5	12.2	446	5 09VZC2	09VZC2 drosophila
8	144	12.2	1209	11 09QX70	09QX70 rattus norv
9	140	11.8	585	12 041935	041935 muria herpe
10	137.5	11.6	358	4 008805	008805 homo sapien
11	136.5	11.5	438	10 039495	039495 cylindrothe
12	136	11.5	1210	11 09EP98	09EP98 mus musculu
13	135.5	11.5	297	4 016038	016038 homo sapien
14	134.5	11.4	379	13 091810	091810 xenopus lae
15	134	11.3	309	5 018751	018751 caenothadl
16	134	11.3	382	4 000599	000599 homo sapien

17	134	11.3	674	10 065672	065672 arabidopsis
18	133.5	11.3	440	4 08TE44	08TE44 homo sapien
19	133.5	11.3	530	11 008934	008934 mus musculu
20	132.5	11.2	322	10 009084	009084 lycopersico
21	132.5	11.2	442	10 039494	039494 cylindrothe
22	132.5	11.2	487	11 08VDN4	08VDN4 rattus norv
23	132.5	11.2	584	16 09FCJ3	09FCJ3 streptomyc
24	131.5	11.1	473	10 039620	039620 chlamydomon
25	131	11.1	903	4 09UPX1	09UPX1 homo sapien
26	130.5	11.0	1480	4 096C04	096C04 homo sapien
27	130.5	11.0	2135	4 043157	043157 homo sapien
28	130.5	11.0	2135	4 09UIV7	09UIV7 homo sapien
29	130	11.0	847	10 09XIB6	09XIB6 arabidopsis
30	130	11.0	3325	12 09IBT9	09IBT9 turkey herp
31	130	11.0	3342	12 09E6N3	09E6N3 turkey herp
32	129.5	11.0	476	5 09XTT6	09XTT6 caenothadl
33	129.5	11.0	1620	11 08V159	08V159 mus musculu
34	129	10.9	440	4 08TF74	08TF74 homo sapien
35	128	10.8	889	16 09F2N5	09F2N5 streptomyc
36	127.5	10.8	551	13 013003	013003 gallus galli
37	127.5	10.8	594	13 09DF69	09DF69 gallus galli
38	127.5	10.8	955	10 094F92	094F92 chlamydomon
39	127.5	10.8	3179	12 08V2A4	08V2A4 human herpe
40	127	10.7	369	10 040692	040692 oryza sativ
41	127	10.7	763	2 09XDH2	09XDH2 mycobacteri
42	126.5	10.7	356	12 09Q0B5	09Q0B5 herpes simp
43	126.5	10.7	734	11 088970	088970 mus musculu
44	126	10.7	562	16 0923X1	0923X1 rhizobium m
45	126	10.7	990	4 09UG03	09UG03 homo sapien

## ALIGNMENTS

RESULT 1  
018735 PRELIMINARY; PRT; 1259 AA.  
ID 018735  
AC 018735;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Erdb-2.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Yokota H.;  
RT "cDNA cloning of erdb-2 from canine mammary gland.";  
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB008451; BAA23127.1;  
DR HSSP; P11362; 1E6K.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR000494; EGFR\_L\_domain.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002174; Furin-like.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR InterPro; IPR004019; YLP\_motif.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF00069; Pkinase; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 2.  
DR Pfam; PF02757; YLP; 2.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00261; Fu; 3.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00018; EF\_HAND; UNKNOWN\_1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW ATP-binding; Transferase; Tyrosine-protein kinase.  
SO SEQUENCE 1259 AA; 137989 MW; E37364D49C4ACD46 CRC64;

Query Match 86.8%; Score 1026.5; DB 6; Length 1259;  
 Best Local Similarity 86.0%; Pred. No. 5.9e-69;  
 Matches 191; Conservative 9; Mismatches 17; Indels 5; Gaps 1;

QY 1 GAGGWHRRSSSTRSGGDLTGLPSESEARSLAEGSGSVFPGDLGMAKG 60  
 DB 1037 GAGGTARRHRSSSTRNGGELTGLPSESEARSLAEGSGSVFPGDLGMAKG 1096  
 QY 61 LQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRRQPPSPREGPL 120  
 DB 1097 LQSLPSDDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRRQPPSPREGPL 1156  
 QY 121 PAARPAATLER-----AKTLSPGKNGVAKDVFAGGAVENPEYLTPQGAAPQPHPPA 175  
 DB 1157 PPSRPAGATLERPKTLSPKTLSPGKNGVAKDVFAGGAVENPEYLTPQGAAPQPHPPA 1216  
 QY 176 FSPAFLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
 DB 1217 FSPAFLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 1258

## RESULT 2

Q8R2X1 PRELIMINARY; PRT; 367 AA.  
 AC Q8R2X1;  
 DT 01-JUN-2002 (TReMBLrel. 21, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Hypothetical 40.2 kDa protein.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC027080; AAH27080.1; -  
 KW Hypothetical protein.  
 SO SEQUENCE 367 AA; 40163 MW; 0BE0395F9E101B0 CRC64;

Query Match 81.3%; Score 961; DB 11; Length 367;  
 Best Local Similarity 82.0%; Pred. No. 1.2e-64;  
 Matches 178; Conservative 8; Mismatches 31; Indels 0; Gaps 0;

QY 1 GAGGWHRRSSSTRSGGDLTGLPSESEARSLAEGSGSVFPGDLGMAKG 60  
 DB 150 GTGTAHRRHRSSSTRSGGDLTGLPSESEARSLAEGSGSVFPGDLGMAKG 209  
 QY 61 LQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRRQPPSPREGPL 120  
 DB 210 LQSLPSDDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRRQPPSPREGPL 269  
 QY 121 PAARPAATLERAKTSLSPGKNGVAKDVFAGGAVENPEYLTPQGAAPQPHPPA 180  
 DB 270 PPIRAPAGATLERPKTLSPKTLSPGKNGVAKDVFAGGAVENPEYLTPQGAAPQPHPPA 329  
 QY 181 DNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217  
 DB 330 DNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 366

## RESULT 3

Q14256 PRELIMINARY; PRT; 165 AA.  
 AC Q14256;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE C-erb B2/neu protein (Fragment).  
 GN C-ERB B2.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86070181; PubMed=2999974;  
 RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,  
 RA Seeburg P.H., Liberman T.A., Schlessinger J., Franke U.,  
 RA Levinson A., Ulrich A.;  
 RT "Tyrosine kinase receptor with extensive homology to EGF receptor  
 RT shares chromosomal location with neu oncogene."  
 RL Science 230:1132-1139(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94000386; PubMed=8104414;  
 RA Sarkar F.H., Ball D.E., Li Y.W., Crisman J.D.;  
 RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)  
 RT gene."  
 RL DNA Cell Biol. 12:611-615(1993).  
 DR EMBL: M95667; AAC37531.1; -  
 FT NON\_TER  
 SO SEQUENCE 165 AA; 17327 MW; A0C113BA308BFA6B CRC64;

Query Match 76.6%; Score 906; DB 4; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-61;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AGSVFPGDGLGMAKAGLQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYV 103  
 DB 1 AGSVFPGDGLGMAKAGLQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYV 60  
 QY 104 NOPDVRRQPPSPREGPLPAARPAATLERAKTSLSPGKNGVAKDVFAGGAVENPEYLTPQ 163  
 DB 61 NOPDVRRQPPSPREGPLPAARPAATLERAKTSLSPGKNGVAKDVFAGGAVENPEYLTPQ 120  
 QY 164 GGAAPQPHPPAPAFNDNLYYWDQDPPERGAPSTFKGTPTAEN 208  
 DB 121 GGAAPQPHPPAPAFNDNLYYWDQDPPERGAPSTFKGTPTAEN 165

## RESULT 4

Q8WYV0 PRELIMINARY; PRT; 412 AA.  
 AC Q8WYV0;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Hypothetical 44.7 kDa protein.  
 GN PP3659.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.O., Qin W.X., Zhao X.T.,  
 RA Wan D.F., Gu J.R.;  
 RT "Novel human cDNA clones with function of inhibiting cancer cell  
 RT growth."  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF318349; AAL55856.1; -  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00069; pkinase.1.  
 DR Pfam: PF02757; YLP\_2.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00219; TyrcKc.1.  
 DR PROSITE: PS00018; EF\_HAND; UNKNOWN\_1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Hypothetical protein.  
 SO SEQUENCE 412 AA; 44702 MW; 034397FF3F27D2BC CRC64;

Query Match 75.7%; Score 894.5; DB 4; Length 412;  
 Best Local Similarity 69.3%; Pred. No. 1.4e-59;  
 Matches 181; Conservative 5; Mismatches 22; Indels 53; Gaps 4;

QY 1 GAGGVNHHNRSSSTRSGGDLTLGLEPSEEPAPSPAPSGAGSDVFDGLGMAKG 60  
 |||||  
 DB 150 GAGGVNHHNRSSSTRSGGDLTLGLEPSEEPAPSPAPSGAGSDVFDGLGMAKG 209  
 |||||

OY 61 LQSLTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNDPVRPOPSPREGPL 120  
 |||||  
 DB 210 LQSLTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNDPVRPOPSPREGPL 269  
 |||||

OY 121 PAARPAGATLERAKTLPSPKNGVAVDFAFGAVENPEYLTPOGGAAPQ----- 170  
 |||||  
 DB 270 PAARPAGATLERAKTLPSPKNGVAVDFAFGAVENPEYLTPOGGAALSPILLPSAQS 329  
 |||||

OY 111 -----HPPPA---ESPADFNLTYWD-QDPPER----- 193  
 |||||

DB 330 TTSITGRTHSGGLHPPAPSKGLRQRTQSTWVWTCQCEPEGOVRRSPDVSSGREGULTS 389  
 |||||

OY 194 -----GAPSTFKGTPTAEN 208  
 |||||

DB 390 AGIKMEGPPTTSKGTCHARN 410  
 |||||

RESULT 5  
 O9W6F6 PRELIMINARY: PRT: 1137 AA.

AC O9W6F6  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-JUN-2002 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Receptor tyrosine kinase (Fragment).  
 CN ERBB4

OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HINDBRAIN;  
 RX MEDLINE=99263203; PubMed=10328884;  
 RA Dixon M., Lumsden A.;  
 RT "Distribution of neurogranin-1 (nrg1) and erbB4 transcripts in  
 embryonic chick hindbrain";  
 RL Mol. Cell. Neurosci. 13:237-258 (1999).  
 DR HSBP; P11362; IFCK.  
 DR InterPro: IPR000494; EGFRL\_domain.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002174; Furlin-like.  
 DR InterPro: IPR001368; TNFR-like.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF00757; Furlin-like; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF01030; Recep\_L\_domain; 1.  
 DR Pfam: PF02757; YLP; 2.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Euk\_kinase; 1.  
 DR SMART: SM00261; FU; 3.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00652; TNFR\_NGFR\_1; UNKNOWN; 1.  
 KW Kinase; Tyrosine-protein kinase.  
 FT NON\_TER  
 SQ SEQUENCE 1137 AA; 127927 MW; 4D616436F87DC4F CRC64;

Query Match 14.0%; Score 165.5; DB 13; Length 1137;  
 Best Local Similarity 26.7%; Pred. No. 0.00024;

Matches 62; Conservative 27; Mismatches 92; Indels 51; Gaps 10;

OY 16 RSGGDLTLGLEPSEEPAPSPAPSGAGSDVFDGLGMAKGLSLPTDPSPL 72  
 |||||  
 DB 899 RDGGYABQGV-PMYRAPGCTIPEAPVAGQATVAFEDTCCNGTLRKQVATLAKEDSST 957  
 |||||

OY 73 QRYSDPTVPLPS-----ETDGYVAPLTCSPQPEYVNDPVRPOPSPREGPLAA-R 124  
 |||||  
 DB 958 QRYSDPTVPLPS-----ETDGYVAPLTCSPQPEYVNDPVRPOPSPREGPLAA-R 1017  
 |||||

OY 125 PAGATLERAKTLPSPKNGVAVDF-----AFGAVENPEYLTPOGGAAPQHPAPAF 176  
 |||||  
 DB 1018 PEYHN-----APNGQPAEDEYVNEPLYLNTFANTLENAEYL-----KNLPEKA 1062  
 |||||

OY 177 SPAPDNLYWMDPPERGA--PSTFKGTP-----AENPEYL 212  
 |||||

DB 1063 KKAFDNPTWNSLPPRSTLQHDYLOEXSTKYFKQNGRIPIVAENPEYL 1114  
 |||||

RESULT 6  
 O8S247 PRELIMINARY: PRT: 420 AA.

AC O8S247  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE RE17165P.  
 CN CG15021.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyridae; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brockstein P., Hong L., Aghayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiliker S.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY071124; AAL48746.1;  
 SQ SEQUENCE 420 AA; 42947 MW; 96D62FFDC9F996E1 CXC64;

Query Match 12.2%; Score 144.5; DB 5; Length 420;  
 Best Local Similarity 28.8%; Pred. No. 0.003;  
 Matches 57; Conservative 11; Mismatches 79; Indels 51; Gaps 10;

OY 34 PRSPLAPSEAGSDVFDGLGMAKGL-----QSLPTHDPSPLQRYSDPTVPLPS 85  
 |||||  
 DB 37 PSVPF-PPPGSGNGIEDSGICGPPAPASAPASVGPQRRPPPPPPQ-----PTPPAPR 90  
 |||||

OY 86 ETDGYVAPLTCSPQPEYVNDPVRPOPSPREGP-LPAARPAGATLERAKTLPSPKNGV 143  
 |||||

DB 91 PS--YGPQTQPPRRPPQPTESA-PAPPPPSGPPQTTPRRPPPPPTESA----- 137  
 |||||

OY 144 VKDVFAGAVENPEYLTPOGGAAPQHPAPSPADFNLTYWMDP-----PERGAPST 199  
 |||||

DB 138 -----PAPPPSYGPPQ--TPPPRPPPPQPTPSAPASVGPQPPAPAPSPSPSP 185  
 |||||

OY 200 FKGTPTAENPEYLGIDVP 217  
 |||||

DB 186 QPG-----PEYLPDP 197  
 |||||

RESULT 7  
 O9VZC2 PRELIMINARY: PRT: 446 AA.

AC O9VZC2  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMURel. 17, Last annotation update)  
 DE CG15021 protein.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:  
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:  
 OC Ephydroidea: Drosophilidae: Drosophila.  
 NCBI\_Taxid=7227;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.C., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Plamkoc C., Baldwin D.,  
 RA Balilew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Chertis J.M., Cawley S., Dahler C., Davenport L.B., Davies P.,  
 RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mohariy C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rainett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slater E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Sykes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003481; AAF47902.1;  
 DR FlyBase: FBgn003534; CG15021.  
 DR Interpro: IPR003882; Pfam: Pf01021.  
 DR Interpro: IPR002965; Pfam: Pf01021.  
 DR PRINTS: PRO1217; PRICHTEXTENS.  
 DR PRINTS: PRO1218; PSTLTEXTENS.  
 SO SEQUENCE 446 AA; 45728 MW; A2C9CE02A07F0542 CRC64;

Query Match 12.2%; Score 144.5; DB 5; Length 446;  
 Best Local Similarity 28.8%; Pred. No. 0.0032;  
 Matches 57; Conservative 11; Mismatches 79; Indels 51; Gaps 10;

QY 34 PRSLASBEGSDVFEGLGMAKAGL-----OSLPHDPSLQRYSEDPVLP 85  
 DB 63 PSVPF-PPSSGNGIEDSGICPPAPAPAPSYGPPQTPPPPPPPPPQ-----PTPAPR 116  
 QY 86 ETGGYVAPLCSQPEVEVNPDPVAPPPSPREPP--LPAPRACATLERAKTISPKNGV 143  
 DB 117 PS--YGPQTPPP 163  
 QY 144 VKDFAFGAVENPEYLTPGGGAAPQHPAPAPAFNDLYYMDQ-----PERGAPST 199  
 DB 164 -----PAPPPSYGPPQ-----TPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 211

QY 200 FKGTAEENPEYGLDVP 217  
 DB 212 QPG-----PEYLPDPQ 223

RESULT 8  
 ID 09QX70 PRELIMINARY; PRT; 1209 AA.  
 AC 09QX70;  
 DT 01-MAY-2000 (TREMURel. 13, Created)  
 DT 01-MAY-2000 (TREMURel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMURel. 21, Last annotation update)  
 DE Epidermal growth factor receptor.  
 GN EGFR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi;  
 OC Mammalia: Eutheria: Rodentia: Sclurognathi: Muridae: Murinae; Rattus.  
 NCBI\_Taxid=10116;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RX MEDLINE=90258808; PubMed=2342466;  
 RA Betch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,  
 RA Earp H.S.;  
 RT "A truncated, secreted form of the epidermal growth factor receptor is  
 encoded by an alternatively spliced transcript in normal rat tissue.";  
 RL Mol. Cell. Biol. 10:2973-2982(1990).  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RA Betch L.A.;  
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FISHER; TISSUE=LIVER;  
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M37394; AAF14008.1;  
 DR HSSP: P11362; IFGK.  
 DR Interpro: IPR000494; EGFR\_L\_domain.  
 DR Interpro: IPR000719; Euk-kinase.  
 DR Interpro: IPR002174; Euk-kinase.  
 DR Interpro: IPR001245; Tyr-kinase.  
 DR Pfam: PF00757; Furin-1like.1.  
 DR Pfam: PF00069; kinase.1.  
 DR Pfam: PF01030; Recept\_L\_domain.2.  
 DR PRINTS: PRO1009; TYRKINASE.  
 DR Prodom: PD000001; Euk-kinase.1.  
 DR SMART: SM00261; FU.3.  
 DR SMART: SM00219; TYRK.1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE: PS00101; PROTEIN\_KINASE\_DOM.1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR.1.  
 KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.  
 SO SEQUENCE 1209 AA; 134891 MW; 96FE7F6C1B7773 CAC64;

Query Match 12.2%; Score 144; DB 11; Length 1209;  
 Best Local Similarity 26.8%; Pred. No. 0.011;  
 Matches 59; Conservative 22; Mismatches 63; Indels 76; Gaps 11;

QY 47 DVFGDLGMAKAGLGLPHDPS-----LQRY 76  
 DB 1010 DVVDADEYELIPQGFNSPSTSRPLSLSSANSNSTVACINRNGSCRYKEDAFQRY 1069  
 QY 77 EDPTVPLPSET--DGVAAPLCSQPEVEVNPDPVAPPPSPREPP--LPAPRACATLERAK 134  
 DB 1070 SDPTSVLTEDNIDTFL-----PVPEYINQ-SVKKPPAGSVONPVYHNQPLHP----- 1116  
 QY 135 TLSPGKNGVAVKDVFAFGAVENPEYLTPGGGAAPQHPAPAPAFNDLYYMDQ----- 188  
 DB 1117 --APGRHLNQN--PHNNAVSNEPEYLNTAQ-----PTCLSSGDSALMIQKQSHQ 1163



QY 189 ----DP-----PERGAPSTFKGPTPAENPEYLGLDVP 217  
 Db 1164 MSLDNPDYQODFFPKAKPNCIFKG-PTAENAEYLRAVP 1202

RESULT 9  
 041935

ID 041935 PRELIMINARY; PRT; 585 AA.  
 AC 041935;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE Hypothetical 60.2 kDa protein.  
 GN GAMMAHV. M6.  
 OS murid herpesvirus 4.  
 OC viruses: dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae.  
 OX NCBI\_TaxID=33708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WMMS;  
 RX MEDLINE=9736649; PubMed=9223479;  
 RA Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,  
 RA Dai Canto A.J., Speck S.H.;  
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus  
 RT 68."  
 RL J. Virol. 71:5894-5904(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WMMS;  
 RA Latreille P., Wamsley P., Waterston R.H.;  
 RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: U97553; AAB6392.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 585 AA; 60160 MW; 85610AAB0C34827D CRC64;

Query Match 11.8%; Score 140; DB 12; Length 585;  
 Best Local Similarity 26.5%; Pred. No. 0.0094;  
 Matches 60; Conservative 8; Mismatches 90; Indels 68; Gaps 10;

QY 28 PSEEARSPAPSE-GAGSVYFGDGLGMAKGLQSLFTHDPFLQRTSDPTVPLPSE 86  
 Db 211 PSMGPDRPPPELPGSPTSPAPSRAGA-----RIPDL-PGLPSMGDPDRPPPP 264  
 QY 87 TDGVYALTCSPOPEYVNO-----PDVRPOPSPSRGP-----LPAARPGAT 129  
 Db 265 ELGSGSP--ISPAISRAGARIPDLGPLPSMGPPRRPPPELPGSPTSPAPSRAGAR 322  
 QY 130 LERAKTSPGKNGVYKDVAFAGAVENPEYLTPOGAAPQ-PHPPPAFSPAFDN----- 182  
 Db 323 IPDL-----PGP-----LPSMGDPDRPPPPPELPGSPTSPAPSR 358  
 QY 183 -----LYYWDODPPERGAPSTFKGPTPAENPEYLGLDVP 217  
 Db 359 AGARIPDLGPLPSMGDPDRPPPPPELPGSPTSPAPSRAGARIP 404

RESULT 10  
 008805

ID 008805 PRELIMINARY; PRT; 358 AA.  
 AC 008805;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PRBL protein (Fragment).  
 GN PRBL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93304421; PubMed=8317492;

RA Azen E.A., Latreille P., Niece R.L.;  
 RT "PRBL gene variants coding for length and null polymorphisms among  
 RT human salivary Ps, Pm, PMS, and Pe proline-rich proteins (PRPs).";  
 RL Am. J. Hum. Genet. 53:264-278(1993).  
 DR EMBL: S62941; AAB27289.1; -  
 FT NON-TER  
 SQ SEQUENCE 358 AA; 35050 MW; DB7F87B8D5EA759E CRC64;

Query Match 11.6%; Score 137.5; DB 4; Length 358;  
 Best Local Similarity 26.6%; Pred. No. 0.0084;  
 Matches 59; Conservative 17; Mismatches 75; Indels 71; Gaps 11;

QY 11 RSSSTRSGGDLTLGLPSEEARSPRLAP---SEG---AGSVFDGLGMAKGLQSL 64  
 Db 52 KSRSPRSPPPK-PQGPFGGNGQGPFPFGKPGQPGGKPKPGQ----- 106  
 QY 65 PTHDSPLOKSEDPYPLPSEIDGYAPLTCSFQPEYVQPDVRPOP----- 113  
 Db 107 ----PPQGDKSQSPRSP-PGKPGQ-----PPQGGNQPGPPPPPKPGFPQGC 153  
 QY 114 SPREGPLAARPGATLERAKTLIS---PGKNGVYKDVAFAGAVENPEYLTPOGAAPQ 169  
 Db 154 NKPGGPPPPPKPGQPPQGDKSQSPRSPPK-----PQGPFGGNGQ 197  
 QY 170 PHPPPAFSPAFDNLYYWDODPPERGA-----PSTFKGTP 204  
 Db 198 GPPPPPKP-----QGPPQGGNRPQGPFPKPKPGQGP 230

RESULT 11  
 039495

ID 039495 PRELIMINARY; PRT; 438 AA.  
 AC 039495;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Alpha 2 frustulin.  
 GN FRU ALPHA 2.  
 OS Cyliotheca frustiformis (Marine diatom).  
 OC Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;  
 OC Bacillariophycidae; Bacillariales; Bacillariaceae; Cyliotheca.  
 OX NCBI\_TaxID=2853;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96314479; PubMed=8706728;  
 RA Kroegeer N., Bergsdorf C., Sumper M.;  
 RT "Frustulins: Domain conversation in a protein family associated with  
 RT diatom cell walls."  
 RL Eur. J. Biochem. 239:259-264(1996).  
 DR EMBL: X99327; CAA67702.1; -  
 SQ SEQUENCE 438 AA; 45947 MW; 3F894F5033634B03 CRC64;

Query Match 11.5%; Score 136.5; DB 10; Length 438;  
 Best Local Similarity 27.8%; Pred. No. 0.012;  
 Matches 57; Conservative 20; Mismatches 55; Indels 73; Gaps 12;

QY 34 PRSPAPSE-----AGSVF-----DGDLGMAKGLQSL----- 63  
 Db 108 PGVPLERCGDCSDSCASSDLFCFPRNDVYVPGRGSGSDSKDYCIKRTADAPGV 167  
 QY 64 ---LPTHDPSPLOKSEDPYPLPSEIDGYAPLTCSPOPEYVQPDVRPOP---PSPRE 117  
 Db 168 APVYPTADPSP-----DPT-PDPSPP-----TPPSPSPVNSPPDPPTPGPTPTP 214  
 QY 118 GPLAARPAQ-----ATLERAKTLISG-----KNGVY---KDV-----FAFGAVE 155  
 Db 215 GPTPASAPSGDPPKATYFPLICGDCDDDPDCEGLICFORDANESVPGSGSSSD 274  
 QY 156 --NPEYLTPOGAAPORHPPPAFSP 178  
 Db 275 RSKTDYCIKENPSTPTPPPPAPAP 299

	RESULT	12
ID	Q9EP98	PRELIMINARY; PRT: 1210 AA.
AC	Q9EP98:	
DT	01-MAR-2001 (TREMBLrel. 16, Created)	
DR	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT	01-JUN-2002 (TREMBLrel. 21, last annotation update)	
DE	Epidermal growth factor receptor isoform 1.	
CN	EGFR.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_Taxid=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C3H/10J, 129/SVJ AND 129/SEVYAC.	
RA	Reiter J.L., Threagde J.L., Eley G.D., Strunk K.E., Danielson A.J.,	
RA	Singclair C.S., Pearisall R.S., Green P.J., Yee D., Lampland A.L.,	
RA	Balsubramanian S., Crossley T.O., Magnuson T.R., James C.D.,	
RA	Malhe N.J.;	
RT	"Comparative genomic sequence analysis and isolation of human and	
RT	mouse alternative Egfr transcripts encoding truncated receptor	
RL	isoforms";	
RL	Submitted (JUN-2000) to the EMBL/Genbank/DBAJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J;	
RA	Reiter J.L., Threagde J.L., Eley G.D., Strunk K.E., Danielson A.J.,	
RA	Scheil C., Pearisall R.S., Green P.J., Yee D., Lampland A.L.,	
RA	Balsubramanian S., Crossley T.O., Magnuson T.R., James C.D.,	
RA	Malhe N.J.;	
RT	"Comparative genomic sequence analysis and isolation of human and	
RT	mouse alternative Egfr transcripts encoding truncated receptor	
RL	isoforms";	
RL	Submitted (JUN-2000) to the EMBL/Genbank/DBAJ databases.	
DR	EMBL; AF275366; AAC28045.1; .	
DR	EMBL; AF275364; AAG28045.1; JOINED.	
DR	EMBL; AF275365; AAG28045.1; JOINED.	
DR	EMBL; AF275367; AAG24386.1; .	
DR	HSSP; P11362; IFGR.	
DR	MGD; MGI:95294; Egfr.	
DR	InterPro: IPR000345; Cytc_heme_bind.	
DR	InterPro: IPR000494; EGFR_Ldomain.	
DR	InterPro: IPR000719; Euk_pkinase.	
DR	InterPro: IPR002174; Furin-like.	
DR	InterPro: IPR002290; Ser_thr_kinase.	
DR	InterPro: IPR001245; Tyr_pkinase.	
DR	Pfam: PF00757; Furin-like; 1.	
DR	Pfam: PF00069; pkinase; 1.	
DR	Pfam: PF01030; Recep_Ldomain; 2.	
DR	PRINTS; PR00109; TYRKINASE.	
DR	Prodom; PD000001; Euk_pkinase; 1.	
DR	SMART; SM00261; FU; 5.	
DR	SMART; SM00220; S.TKC; 1.	
DR	SMART; SM00219; TyrcK; 1.	
DR	PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.	
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.	
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.	
KM	ATP-binding; Receptor; Transferase.	
SEQ	SEQUENCE 1210 AA; 134840 MW; 62CD021C9DE32E18 CRC64;	
Query Match	11.5%; Score 136; DB 11; Length 1210;	
Best Local Similarity	26.4%; Pred. NO. 0.043;	
Matches	58; Conservative 23; Mismatches 63; Indels 76; Gaps 11	
OY	47 DYFDGDLGAKAGKGLSLFTHDPSP-----LQRYS 76	
DB	1011 DVAVDDEVLTLPQGFGFNSTSTRTPLLSLSATSNNSTVACINRNCSCKVKEDAFLOKRS 1070	
OY	77 EDPTVPPLSET--DGVARVLTCSPQPEYVNODVVRQPPSPREGPLPARPAAGATLERAK 134	
DB	1071 SPTPGCAVTEDNIDDAFL-----VPEPYNO-SVPRKPAGGSVONPVYHNQPLHP----- 1117	

Query Match	Best Local Similarity	Score	DB	Length
Matches 64; Conservative 17; Mismatches 76; Indels 94; Gaps 13.	25.5%; Pred. Num. 0.0036;	11.5%;	297 AA;	1079
11 RSSSTRSGGGDITLLESEEEAPRSLAP---SEG---AGSDVFDGDLGMAKGLQSL 64	52 KRSRSRSPGPK-PGGRPPGQGNQPGGRPPGPKPGGRPPGQGNKPGGRPPGPKPGQ--- 106	65 PTHDPSPLQKRSSEDPVLPSETDGVAPLTCSPQPEYVNPDPVRPQPS-----PR--- 116	107 ----PPQGDKSQSPRSP-PGKPGQ-----PPQGNQGNQPGGRPPGPKPGQPGQGG 153	117 ----EERPLAARAGATLERAKTLS----PGKNGVYKDVAFGAGVANEPEYLTPOGGAAPQ 169
154 NRPQSPRPPGPKQGGPPQGDKSQSPQSPGK-----PGGRPPGQGNQ 197	170 PHPPAFSAPFENLYMDODDPERG-----ALVSTPK 201	198 GPPRPPGPK-----QGPRPPGQGNKPGGRPPGPKPGGRPAQSGSKSQAARAPGPKRQ 248	202 GTPTAE--NPE 210	249 GPPOEGNNPQ 259
RESULT 14	PRELIMINARY; PRT; 379 AA.			
091810	091810	091810	091810	091810
01-NOV-1996 (Tremblrel. 01, Created)	01-NOV-1996 (Tremblrel. 01, Last sequence update)	01-NOV-1996 (Tremblrel. 01, Last sequence update)	01-DEC-2001 (Tremblrel. 19, Last annotation update)	01-JUN-2002 (Tremblrel. 21, Last annotation update)
Proline rich protein.	Xenopus laevis (African clawed frog).	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Xenopodine; Xenopus.	NCBI_TaxID=8355;	NCBI_TaxID=8355;

RP SEQUENCE FROM N.A.  
 RC TISSUE=OOCYTE;  
 RA Nishimatsu S.I., Satoshii, Oda, Naoto, Ueno;  
 RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X68249; CAA48321.1; .  
 DR InterPro: IPR003124; WH2; 1.  
 DR Pfam: PF02205; WH2; 1.  
 DR PRINTS: PRO1574; TUBBYPROTEIN.  
 DR SMART: SM00246; WH2; 1.  
 SO SEQUENCE 379 AA; 38859 MW; 84DC0FB24F971AAD CRC64;

Query Match 11.4%; Score 134.5; DB 13; Length 379;  
 Best Local Similarity 28.0%; Pred. No. 0.015;  
 Matches 59; Conservative 16; Mismatches 97; Indels 39; Gaps 9;

OY 2 AGGWNHHNRSSSTRSGGGDLTGLPESEEARPSRLAPSESG---AGSDVDFDGLGMGA 58  
 DB 134 SGGRRHDDSDGNSRSPPEVGRHAPSLPDLSPRPSSTSGMKHSSAPPPPPGRROA 193  
 OY 59 KGLSLPTPHDPSPLQRYSEDPVPLPSEDTGVAPLTCSPOPEYVQPDVPRQPPSP--- 115  
 DB 194 -GAPRAPQGNKPKYR--EKPLPPTPGHAPAPAPVKKPPSPINSRPSAHSQPPPPPY 250  
 OY 116 -----REGPLPAAPRGATLERAKTISPGKNGYKDVFAFGAVENPE--YLTPOGGA 166  
 DB 251 RQPTSLNGPPSPINEPAPELPQRHNSLHRTAGFVR-----GLAPPPQSVHLSP- 304  
 OY 167 APQHPAPAFSPAFDNLTYWDDPPERGAP 197  
 DB 305 RP---PPA-----RDPPRGGAAP 320

## RESULT 15

O18751 PRELIMINARY: PRT: 309 AA.  
 AC O18751.  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)  
 DE C50F7.5 protein.  
 GN C50F7.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE-94150718; PubMed-7906398;  
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA Johnson, D., Steillyes L.;  
 RT "The sequence of C. elegans cosmid C50F7.";  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U41557; AAA83301.1; .  
 SO SEQUENCE 309 AA; 31085 MW; 6171574A991696D7 CRC64;

Query Match 11.3%; Score 134; DB 5; Length 309;  
 Best Local Similarity 24.2%; Pred. No. 0.013;  
 Matches 54; Conservative 25; Mismatches 74; Indels 70; Gaps 10;

OY 7 HHRHRSSTRSGGGDLTGLPESEEARPSRLAPSESGSDVDFDGLGMGAAGLQSLPT 66  
 DB 31 HHHHKTAPRTSRGIATTTFAFTSSDL---PIAGSSSA----- 65  
 OY 67 HDPSPLQRYSEDP-----TVPLPSE--TDGYVAPLTCSPOPEYVQPDVPRQPPSPRGP 119  
 DB 66 -----PVIASSADPILPITSVVPQPSNEPSPGTVAP---SDEPSPSGPPS--PGVNPSEDP 116  
 OY 120 LPAAPAGATLERAKTISPGKNGYKDVFAFGAVENPEYLTPOGGAPOP-----HPP 174  
 DB 117 QPSGPPSPGPVDPSEDPQ-----SVEPSEDPQPSGPPSPGPVDPSEDPQ 162  
 OY 175 AFSAPFDNLTYWDDPPERGAPPTFTKGTPTAENPEYGLDVP 217  
 DB 163 SVEPSED-----PQSGPPSPGPVDP--SEDPQPSGSSSP 195

Search completed: April 28, 2003, 13:41:41  
 Job time : 20.515 secs

**THIS PAGE BLANK (USPTO)**